

General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

NASA CR-156805

JANUARY 1978

**Geosynchronous Microwave
Atmospheric Sounding
Radiometer (MASR)
Feasibility Study**

Final Report

Contract NAS 5-24082

NASA Technical Officer
J Shiue

Hughes Program Manager
F E Goodwin

This Final Report Prepared By
F E Goodwin
M S Hersman
M Luning



**Volume II
Radiometer Receiver Feasibility**

HUGHES

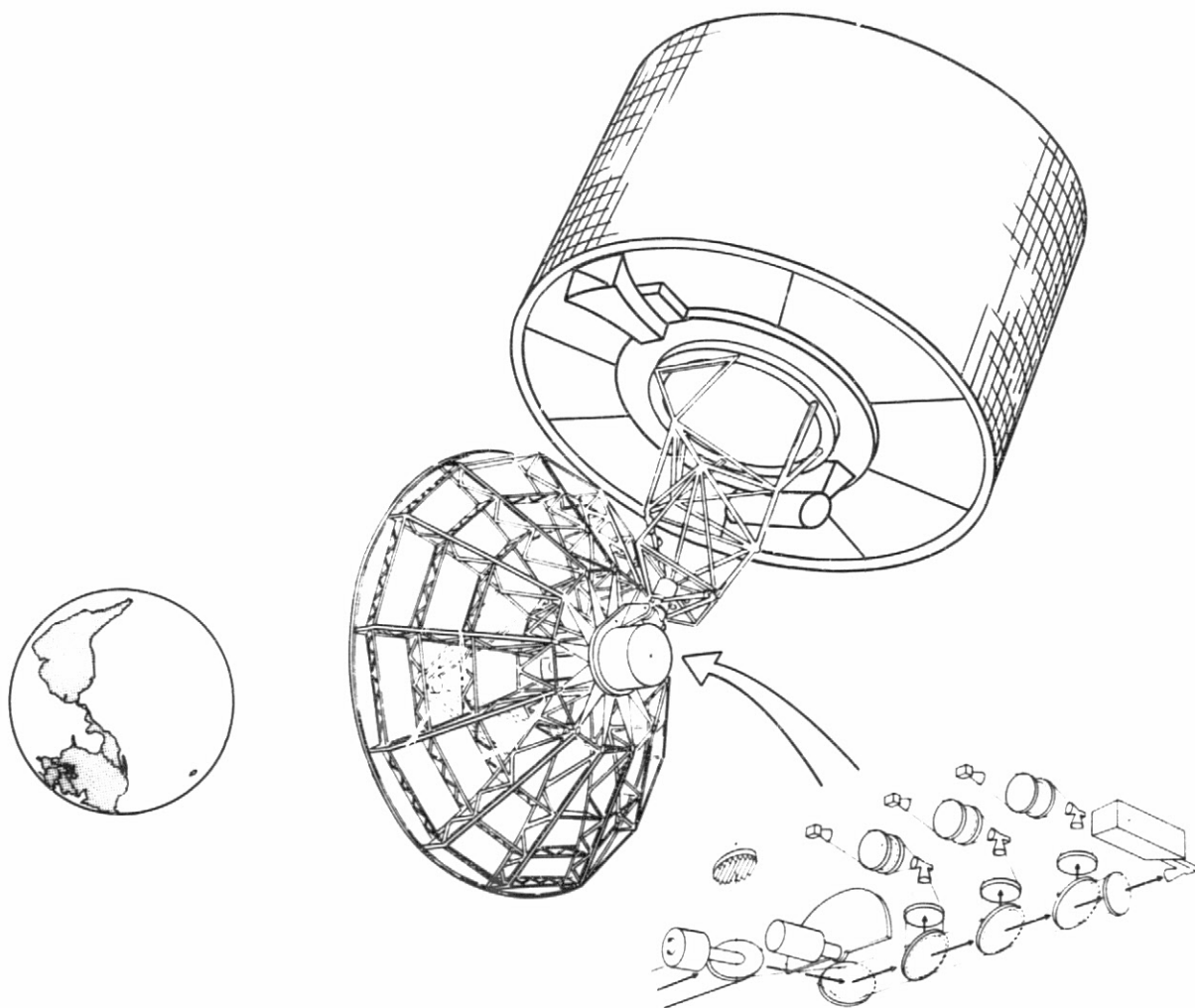
HUGHES AIRCRAFT COMPANY
SPACE AND COMMUNICATIONS GROUP

Hughes Ref No. D8647 • SCG 70532R

(NASA-CR-156805) GEOSYNCHRONOUS MICROWAVE
ATMOSPHERIC SOUNDING RADIOMETER (MASR)
FEASIBILITY STUDY. VOLUME 2: RADIOMETER
RECEIVER FEASIBILITY Final Report (Hughes
Aircraft Co.) 214 p HC A10/ME A01 CSCL 04A G3/46

N78-30749

Unclas
29090



JANUARY 1978

Geosynchronous Microwave Atmospheric Sounding Radiometer (MASR) Feasibility Study

Final Report

Contract NAS 5-24082

NASA Technical Officer
J Shiue

Hughes Program Manager
F E Goodwin

This Final Report Prepared By
F E Goodwin
M S Hersman
M Luming

Volume II Radiometer Receiver Feasibility

HUGHES

HUGHES AIRCRAFT COMPANY
SPACE AND COMMUNICATIONS GROUP

Hughes Ref No. D8647 • SCG 70532R

CONTENTS

	<u>Page</u>
1. INTRODUCTION AND SUMMARY	1-1
1.1 Introduction	1-2
1.2 Summary	1-6
1.3 Recommended Radiometer Configuration	1-14
2. RECOMMENDED RADIOMETER CONFIGURATION	2-1
2.1 Performance Estimates	2-2
2.2 Radiometer Mechanical Subsystem	2-6
2.3 MASR Weight and Power Requirements	2-10
3. DESIGN ANALYSIS	3-1
3.1 MASR Concept	3-2
3.2 Noise Temperature	3-4
3.3 Radiometer ΔT Performance Parametric Tradeoffs	3-12
3.4 Reference Averaging	3-16
3.5 Radiometer Calibration	3-22
3.6 Quasi-Optical Design Formalism	3-28
3.7 Fabry-Perot Filter Design Considerations	3-36
4. RECEIVER DESIGN	4-1
4.1 Mixer Design Alternatives	4-4
4.2 Mixer Diodes	4-16
4.3 Local Oscillator Sources	4-20
4.4 Quasi-Optical Local Oscillator Filters	4-26
4.5 IF Design	4-36
4.6 Quasi-Optical Diplexers	4-40
5. ANTENNA INTERFACE	5-1
5.1 Illumination Patterns	5-4
5.2 MASR Antenna and Feed Geometry	5-6
5.3 Quasi-Optical Quadraplexing	5-10
5.4 Antenna Calibration	5-12
5.5 Systematic Errors	5-20
APPENDICES	
A MASR Concept	A-1
B Reference Averaging in Presence of Gain Fluctuations	B-1
C Laser Beams and Resonators	C-1
D Properties of Materials from 100 to 200 GHz	D-1

1. INTRODUCTION AND SUMMARY

1.1 INTRODUCTION

1-2

1.2 SUMMARY

1-6

1.3 RECOMMENDED RADIOMETER CONFIGURATION

1-14

1. Introduction and Summary

1.1 INTRODUCTION

The geosynchronous microwave atmospheric sounding radiometer (MASR) concept has great merit in meteorological applications. This study outlines an instrument design which meets needed system performance goals, and shows that the instrument technology is feasible.

The microwave atmospheric sounding radiometer is a four-band, millimeter wave radiometer designed to accomplish three-dimensional atmospheric temperature and relative humidity mapping from geosynchronous orbit. The MASR mission collects data to aid in the observation and prediction of severe storms. The geosynchronous orbit allows the continuous atmospheric measurement needed to resolve mesoscale dynamics. The instrument may operate in conjunction with multispectral imaging and infrared sounding sensors, and will be particularly valuable in collecting data in overcast regions. The anticipated ability to predict severe thunderstorms and larger cyclonic storms from MASR data adds to the economic and social benefit.

Angular resolution of MASR is limited by the beamwidth of a shuttle launched antenna (4.4 meter diameter) at the 118 and 183 GHz oxygen and water vapor sounding frequencies. Temperature resolution is a function of system noise temperature and integration time. Resolution element size and required integration time combine to limit the mapping coverage rate. From geosynchronous orbit, the 4.4 meter antenna provides nadir earth surface resolution of about 20 km at 183 GHz. The antenna is mechanically scanned in raster fashion to cover an area with contiguous resolution cells. The nominal dwell time per cell is 1 second at 183 GHz, allowing an area 800 by 800 km to be mapped in about 30 minutes.

Temperature sounding is achieved by obtaining a profile of the thermal radiance of the atmosphere in the frequency region of the 118 GHz oxygen absorption line. There are 11 double sideband (DSB) channels centered at the peak of this line and located various distances from the line center. Each channel will receive radiation predominately from a given altitude layer of the atmosphere, and the atmospheric temperature profile can be obtained from the 11 measured brightness temperatures using an inversion algorithm.

Similarly, there are six DSB channels centered at the 183 GHz water vapor line to determine the humidity profile. In addition, there are two window bands at 104 and 140 GHz to correct the measurements at 118 and 183 GHz for ground emissions (see Figures 1 and 2).

A 4.4 meter symmetric Cassegrain antenna brings all four frequency bands to a common focus. The radiometer receiver package is designed to demultiplex the four bands through the use of quasi-optical filters and to direct the individual beams to four separate feed horns, followed by four separate receivers. The 118 and 183 GHz receivers have very wide band channelized IF systems for the purpose of profiling the microwave spectrum of the atmosphere at these frequencies. Window bands measure the temperature of the earth's surface and therefore, each requires only one IF channel. A common chopper wheel and calibration load serve all four bands. Synchronous detection and other signal processing are achieved through a single onboard digital processor. Figure 3 is a simplified block diagram of the radiometer receiver subsystem.

BAND	BANDWIDTH (IF)	CHANNELS	FUNCTION
183 GHz	10 GHz	6	WATER VAPOR ABSORPTION
140	1	1	WINDOW
118	4.1	11	OXYGEN ABSORPTION
104	1	1	WINDOW

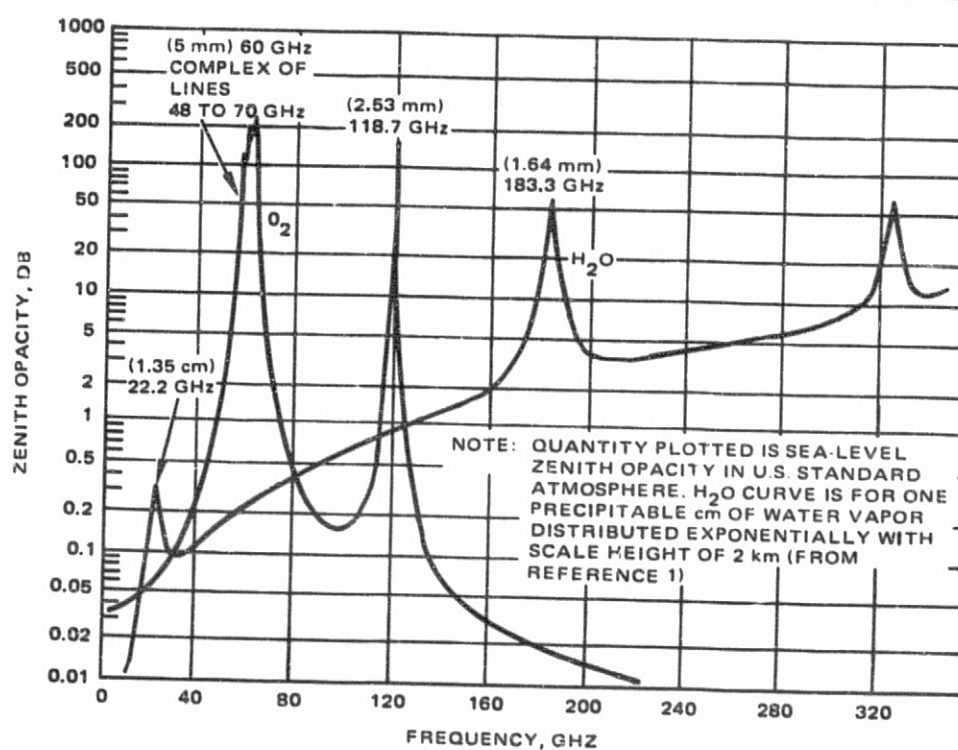


FIGURE 1. MICROWAVE ABSORPTION RESONANCES OF O_2 AND H_2O AVAILABLE FOR SOUNDING

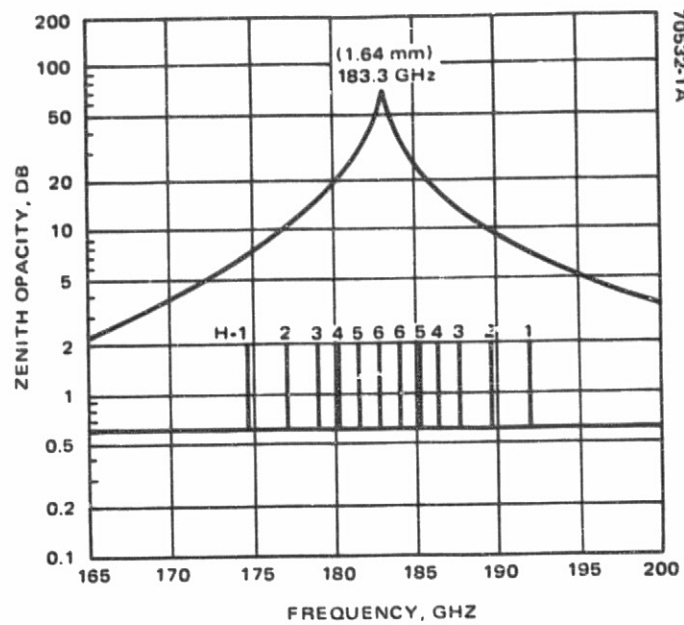


FIGURE 2. H₂O SOUNDING CHANNELS H-1 TO H-6

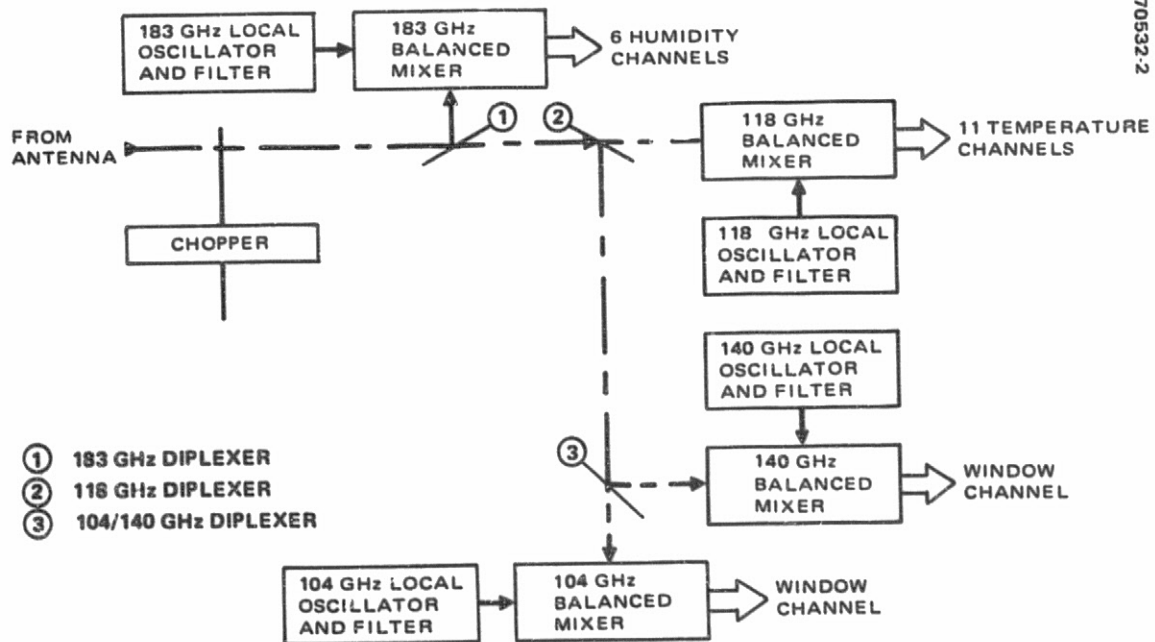


FIGURE 3. RADIOMETER RECEIVER SUBSYSTEM SIMPLIFIED BLOCK DIAGRAM

Millimeter wave technology status offers several approaches to MASR, each having "softness" but currently under rapid development. A preferred technical approach is difficult to recommend at present, but several technical approaches can meet the MASR system requirements. The present study effort has selected a tentative configuration which appears to have least technical risk. Technology advances in alternate approaches could change the selection but the present choice serves as a model to estimate weight, power, size, and cost of the flight instrument.

This report addresses the tradeoff and design study of the MASR radiometer. Volume III gives details of the 4.4 meter antenna subsystem.

The authors are grateful to the following for their assistance in preparing this final report: D. H. Staelin, Massachusetts Institute of Technology; D. N. Held and A. R. Kerr, NASA-Goddard Institute of Space Sciences; T. S. Chu, Bell Telephone Laboratory; J. M. Schuchardt, J. B. Langley, and J. A. Stratigos, Georgia Institute of Technology; J. Payne, B. L. Ulich, National Radio Astronomy Observatory; and J. Shiue, L. King, and T. Dod, NASA Goddard Space Flight Center. Within Hughes, special thanks are given to the Technical Coordination Committee: T. A. Midford, K. L. Brinkman, P. Schwartz, W. H. Kummer, and Chairman R. Graves. Also, the technical assistance of P. Bernues and K. Weller proved to be invaluable.

Specific reference material used in the preparation of this final report is cited in each technical section. General references appear below.

References

1. Proposal for Geosynchronous Microwave Atmospheric Sounding Radiometer, Hughes Ref. 76(41)-22923/D8647,
2. NASA/GSFC RFP NAS5-92451-254, Statement of Work for Geosynchronous Atmospheric Sounding Radiometer Feasibility Study,
3. D. H. Staelin, "Feasibility Study for a Microwave Spectrometer for Meteorological Observations from Synchronous Satellites, "Quarterly Progress Reports 1 and 2, NASA Contract NAS5-23677, 7 July 1977.
4. J. M. Schuchardt, et. al., "94/183 GHz Aircraft Radiometer for Project Storm Fury, " Design Study Report, NASA Contract NAS5-23710, June 1977.
5. J. M. Schuchardt, "Dual Channel Millimeter Wave Radiometer for Convair 990 Flight Tests, " Presentation material, April 1977.
6. F. E. Goodwin, "Microwave Atmospheric Sounding Radiometer, " Presentation material, Contract NAS5-24082, 9-10 August 1977.

1.2 SUMMARY

Design analyses and highlights of technical tradeoffs surrounding a radiometer configuration are described in this summary. Estimated ΔT 's will meet anticipated MASR needs.

MASR study objectives are to make technical tradeoffs, determine feasibility, and isolate key technology areas. These objectives were achieved with in-depth technology surveys, concluding with a receiver hardware definition of sufficient detail to allow weight and power estimates and the preparation of a preliminary specification.

The study revealed the importance of quasi-optical techniques to reduce feed, diplexer circuit, and local oscillator (LO) filter losses. It also identified the mixer as the "softest" technology, including the mixer configuration, mixer diodes, mixer-to-IF impedance match, mixer bandwidth, and mixer local oscillator. While the best state of the art mixer performance was reported to be sufficient to meet MASR requirements, typical performance is in general marginal. Conclusions are that MASR is feasible, ΔT (radiometer temperature resolution) requirements generally can be achieved, quasi-optical design techniques are recommended, and mixer development methodology will pace the program.

This technical report consists of sections on Design Analysis, Receiver Design, and Antenna Interface; appendices deal with the MASR concept, and optical design formalism, reference averaging, and properties of materials. The following paragraphs summarize the highlights of each section.

Design Analysis. Radiometer temperature resolution is commonly expressed as the minimum temperature difference which can be detected

$$\Delta T = KT_S / \sqrt{B\tau}$$

where

T_S = system noise temperature

B = IF bandwidth

τ = integration time

From system analysis of the radiometer, it was determined that the effective sensitivity factor K of the radiometer (normally $K = 2$ for a Dicke switched system) could be reduced to a value approaching 1.4, depending on the gain stability of the system, by using the technique of reference averaging. The improvement in temperature resolution results by integrating the reference channel over a number of pixel periods, where a pixel period equals the integration time of the signal channel. With practical limitations, improvement factor offered by reference averaging is about equivalent to that achievable with a double Dicke (Graham's receiver) system, but does not require the complexity of two receivers.

It is recommended that calibration of the radiometer receiver be performed at the antenna feed with a movable reflector that periodically switches the receiver input to a reference hot load and to cold space. The calibration accuracy expected is of the order of $\pm 1^\circ\text{K}$ in all channels. The radiometer system calibration, including the main antenna, must be performed in the far field of the main antenna. Possible methods of achieving total radiometer system calibration are discussed in Topic 5.4.

It is recommended that quasi-optical filters be used both for LO filters and for frequency multiplexing in MASR. Analytical techniques commonly used for optical system design were found to be advantageous; ray matrices and Gaussian beam formalism provide an extremely versatile design approach for quasi-optics. Unconditional stability of the beam optical system assures negligible loss due to diffraction and radiation from these open structures. Appendix B provides the fundamental reference work of Kogelnik and Li which details the mathematical formalism; Topic 3.7 summarizes the work and illustrates its applications to quasi-optical systems.

Appendix C presents an analysis that determines the effects of gain variation errors in a reference averaging radiometer and the maximum improvement achievable. This analysis shows that for typical gain stability the K value can be reduced from $K = 2$ (Dicke case with square wave modulation and demodulation) to $K = 1.4$. The analysis is summarized and results presented in Topic 3.4.

A theoretical treatment of reference averaging was conducted which includes general considerations of chopping frequency, asymmetrical chopping, and chopper transition effects. Because this work is highly theoretical and lengthy, it is included in Volume IV of this report.

Another theoretical treatment was developed of chopper frequency selection in a reference averaging Dicke radiometer. One of the more significant results is that second detector low frequency noise ($1/f$), which is suppressed in a conventional Dicke radiometer by proper selection of chopper frequency, cannot be suppressed in the same way in a reference averaged system. Rather, the reference averaged system has a floor on the detector noise feedthrough. Because this analysis is highly theoretical and lengthy, it is also included in Volume IV.

Receiver Design. A tentative receiver subsystem block diagram is shown in Figure 1. The basic four-band radiometer receiver consists of a quasi-optical quadruplexer which separates the bands and directs the beams to separate feed horns and mixers. The 118 and 183 GHz mixers are matched into IF networks that provide IF channel separation for the 11 temperature channels and six water vapor channels. The window bands at 104 and 140 GHz are amplified after mixing over nominal IF bandwidths of about 1 GHz and detected. A signal processing and control subsystem manages all functions, including the antenna drive, calibration control, chopper synchronization, and synchronous detection of the 19 radiometer channels. In addition, the signal processing and control subsystem provides the capability to perform reference averaging to enhance the performance of the radiometer.

Radiometer temperature resolution or ΔT performance is a direct function of system noise temperature. To facilitate the following discussion, the contributing

factors are given in Table 1, together with example values for parameters for each radiometer band and a computation of its respective system noise temperature.

Mixer Tradeoff. The heart of the radiometer receiver is the low noise mixer. Reported mixer conversion loss and noise temperature data are scattered over a wide range of values for the frequencies of interest. It is obvious that consistent mixer performance will be achieved only with great effort. A recommended approach is to begin mixer development early and to build a selected inventory of mixers suitable for MASR.

The quality of mixer performance is conventionally expressed in terms of single-sideband noise temperature T_{MSSB} and mixer conversion loss L_M . Mixer performance is affected by the quality of the mixer diode material, precision contacting, microwave circuit design, and impedance match for both the signal input and IF output circuits. Further, it is also a function of the quality and level of local oscillator signal. Of all the factors influencing the radiometer system noise temperature, T_{MSSB} is the term over which there is least control. State of the art values of T_{MSSB} reported in the literature vary over a wide range. Figure 2 shows a number of demonstrated values. The solid line shows the probably minimum achievable values expected, and also represents approximate MASR needs for the more difficult bands.

Balanced mixers have the desirable property of canceling LO noise and are recommended for MASR to provide a degree of immunity to the noise generated by solid state LO sources such as IMPATTs. Waveguide balanced mixers are the recommended baseline design. However, conventional balanced mixers at these frequencies are difficult to build. Subharmonically pumped mixers, a form of balanced mixer, also have LO noise cancellation properties. Further, these devices have the advantage of using a local oscillator at half the signal frequency.

Quasi-Optical Balanced Mixer. This study has produced an alternate approach to achieving the noise cancellation effects of a balanced mixer with two single-ended mixers in a quasi-optical assembly (Figure 3). The signal from the antenna is diplexed with a wire grid polarizer 1 and the 183 GHz band is reflected by 90°. The signal polarization is upward (not shown in figure) while the local oscillator polarization is in the plane (see Figure 3). Both signal and local oscillator encounter a rotation plane of 45° and are split by a second wire grid polarizer 2 and directed to separate single-ended mixer mounts. The signal and LO fields are in phase in one mixer and out of phase in the other. Their electrical outputs combine in such a way as to cancel the dc and noise terms while the IF terms add in phase. This approach was determined to be unique and was reported as a new technology item.

Local Oscillator Sources. It was determined that IMPATT and Gunn solid state sources are available for use as local oscillators. IMPATTs produce sufficient power in all four bands but are noisy and must be filtered with about 40 dB of filtering over each band of frequencies, which for the 183 GHz humidity band is ± 10 GHz from the center frequency. The basic approach to accommodating the IMPATT is to utilize balanced mixers, that provide about 20 dB of immunity to the noise, and Fabry Perot quasi-optical filters that provide at least another 25 dB of filtering. The combined effects of the balanced mixer and quasi-optical filter provide sufficient margin to assure that a good low noise mixer is achieved.

BAND	BANDWIDTH (IF)	CHANNELS	FUNCTION
183 GHz	10 GHz	6	WATER VAPOR ABSORPTION
140	1	1	WINDOW
118	4.1	11	OXYGEN ABSORPTION
104	1	1	WINDOW

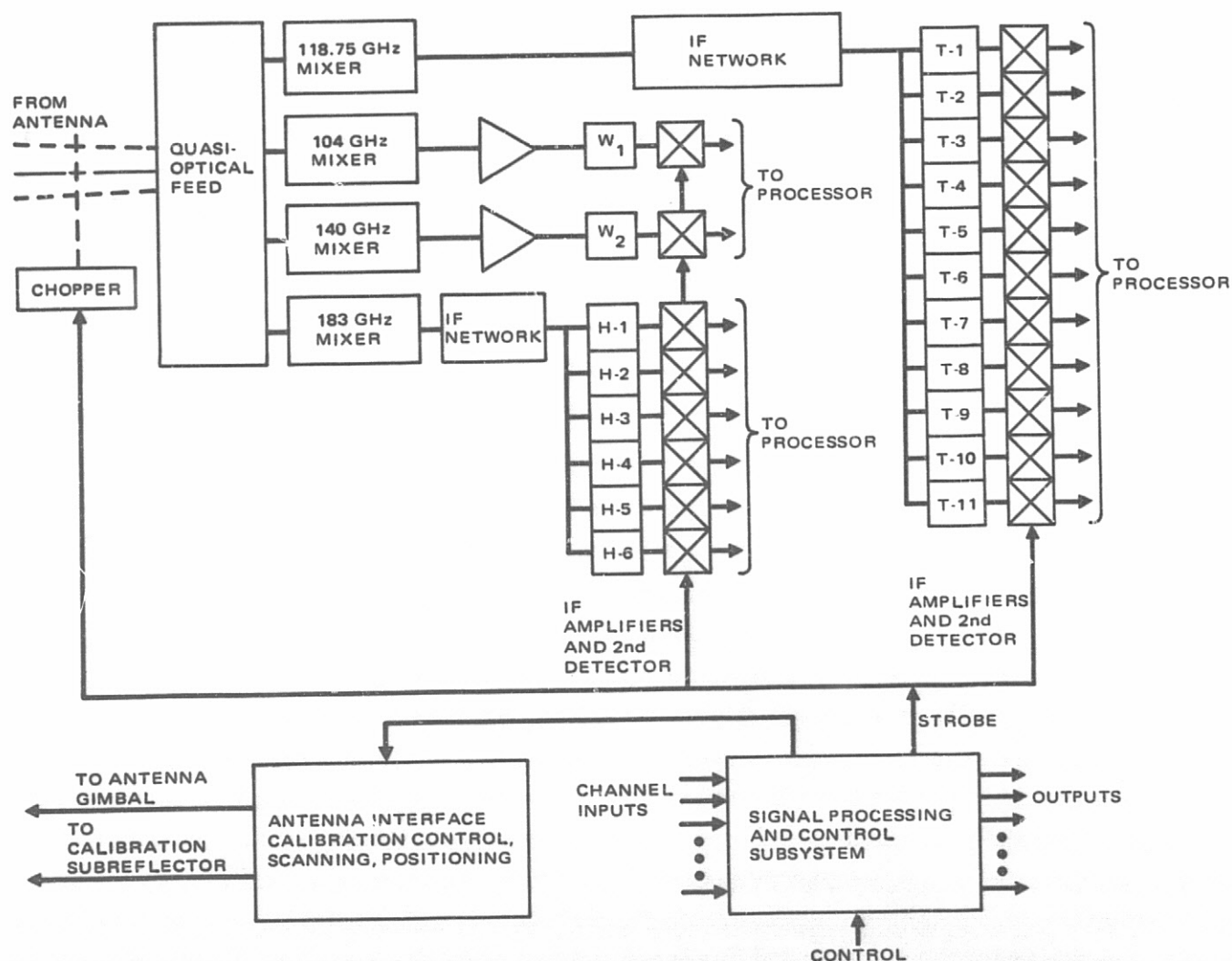


FIGURE 1. FOUR CHANNEL MICROWAVE H₂O AND O₂ SOUNDING RADIOMETER TENTATIVE BLOCK DIAGRAM

TABLE 1. CONTRIBUTORS TO T_{SYSTEM}

$$T_{\text{SYSTEM}} = T_A + (L_F - 1) T_O + L_F T_R$$

$$T_R = \frac{1}{2} (T_{\text{MSSB}} + L_M T_{\text{IF}})$$

T_R = double sideband receiver temperature
 T_A = antenna temperature
 L_F = feed loss
 T_O = feed temperature
 L_M = mixer conversion loss
 T_{IF} = IF noise temperature.
 F_{IF} = IF noise figure

System	183 GHz	140 GHz	118 GHz	104 GHz
$T_A, ^\circ\text{K}$	250	250	250	250
L_M, dB	8.5	6.5	5.0	5.0
F_{IF}, dB	5.0	3.0	3.5	3.0
$T_{\text{MSSB}}, ^\circ\text{K}$	2100	1100	600	600
$L_M T_{\text{IF}}, ^\circ\text{K}$	4438	1288	1135	910
$T_R, ^\circ\text{K}$	3269	1169	867	755
L_F, dB	0.5	1.0	1.0	1.0
$(L_F - 1) T_O, ^\circ\text{K}$	35	70	70	70
$T_S, ^\circ\text{K}$	3952	1823	1412	1270

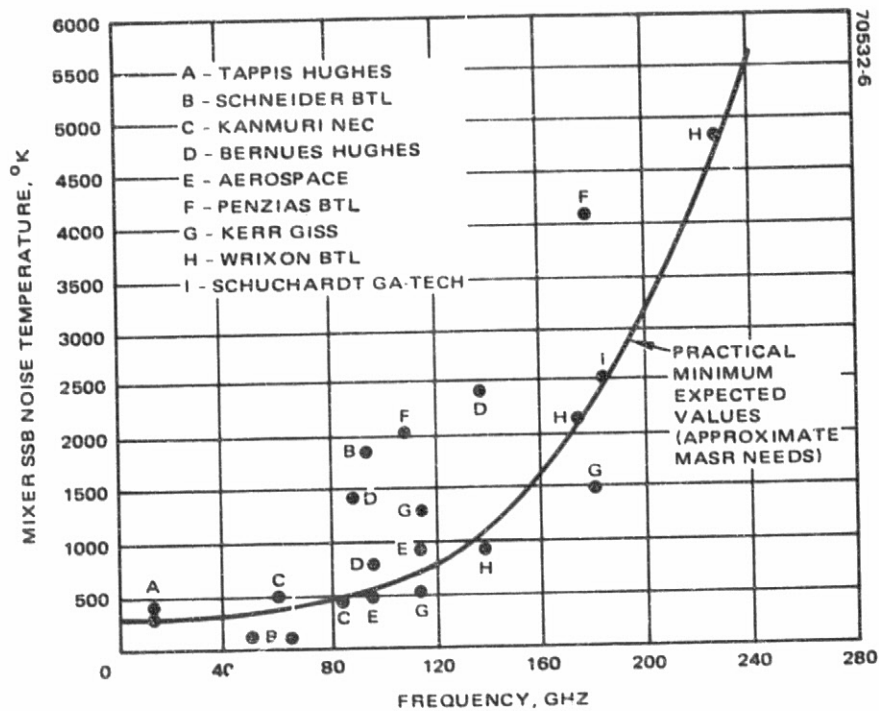


FIGURE 2. DEMONSTRATED LOW NOISE MIXER PERFORMANCE

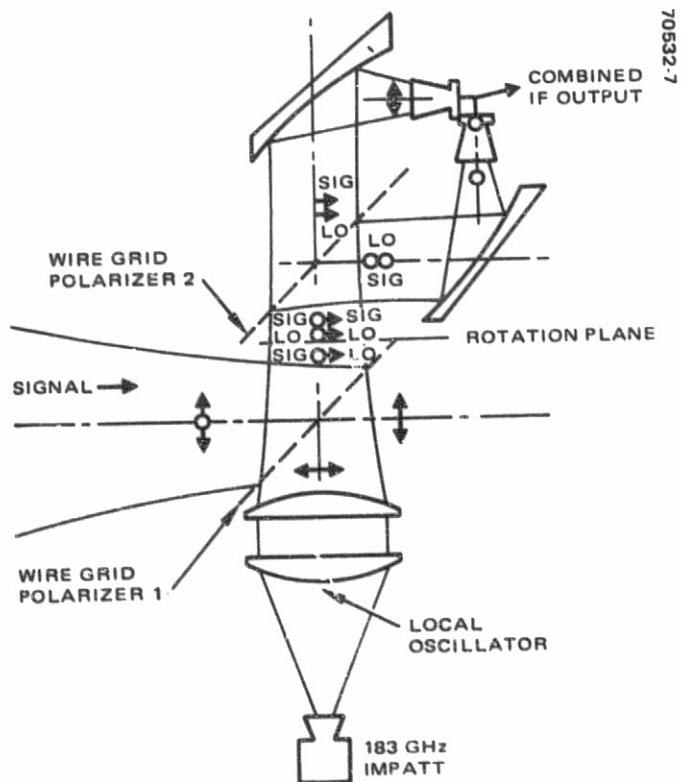


FIGURE 3. QUASI-OPTICAL BALANCED MIXER

The Gunn oscillator (or the transferred electron oscillator) is somewhat quieter than an IMPATT, but has an upper frequency limit near 100 GHz. With noise characteristics lower than that of an IMPATT and adequate power below 100 GHz, the Gunn oscillator is an attractive alternative to the IMPATT when used with a doubler or with a subharmonic mixer.

Local Oscillator Filters. Because of relatively high unloaded Q, Fabry-Perot quasi-optical filters have fundamental advantages over waveguide filters for selectivity and suppression of LO noise in the IF sidebands. Of the two types demonstrated, the parallel plate Fabry-Perot has the advantage of arbitrarily large rejection bands, while the ring configuration offers a unique LO injection scheme. Because of the wide IF bandwidth requirements however, the recommended filter for MASR local oscillators is the parallel plate configuration. Specific parallel plate filter designs were developed for each of the MASR bands. The most difficult filter is that for the 118 GHz band which needs a loaded Q of 6000 to accommodate a minimum IF frequency of only 20 MHz. For a 118 GHz minimum IF frequency of 50 MHz instead of 20 MHz, a loaded Q of 2400 would be adequate, resulting in LO filter losses of 4.1 dB instead of 8.1 dB. The 118 GHz LO power is marginal with 8.1 dB filter loss, and raising the minimum 118 GHz IF would eliminate this problem.

IF Design. Difficult requirements for MASR are the 0.6 to 10 GHz IF network for the H₂O band and the 20 to 4000 MHz IF network for the O₂ band. Two approaches were analyzed; one utilizes available components and techniques while the other assumes the development of contiguous filter banks. The analyses show that while both approaches are feasible and technically suitable, the conventional approach requires less development. In either case, the ultimate performance of the IF network will be determined by how well the mixer-to-IF impedance match is maintained over the required bandwidth.

Antenna Interface. The antenna interface consists of an RF interface, thermal/mechanical interface, and calibration interface. In a typical antenna, the RF interface is defined by a waveguide that connects the receiver to the antenna feed. In the recommended quasi-optical design, the RF interface is defined by the complex beam parameters for the four bands. The illumination patterns for the individual bands are generated separately and combined in the quasi-optical quadruplexer.

The calibration interface deserves special comment. The calibration procedure for the radiometer receiver is self-contained. A positionable mirror at the antenna feed allows the receiver to look at the antenna, cold space, or a calibrated reference "hot" load. This two-point receiver calibration sequence could be performed in a few seconds and repeated often, allowing the receiver to measure antenna temperature T_A to an accuracy of about 1°K.

Knowing T_A accurately is only one step in obtaining mesoscale temperature data since the value of T_A is a function of the antenna beam efficiency, sidelobes, field brightness, and ohmic losses. The ultimate calibration of the actual target temperature against T_A is a difficult task which must be carefully performed in space. Topic 5.4 is devoted to the discussion of antenna calibration.

Quasi-Optical Antenna Feed. The tradeoff study shows that feed, diplexing, and LO filter losses could be significantly reduced by using quasi-optical components rather than waveguide components. Typically, a ten-fold reduction in losses is

illustrated in the radiometer point design. Appropriately, optical design techniques are found to be powerful aids in achieving low loss designs of quasi-optical devices and systems. A novel quasi-optical feed was disclosed (Figure 4) which separates the 183 GHz channel with a wire grid polarizer (element 1) and separates the 118 GHz channel from the window channels with either a Fabry-Perot or resonant grid optical filter (element 2). The antenna feed point is imaged at each of the four horns. Typical feed losses, including 0.25 dB for each horn, are given in Figure 4.

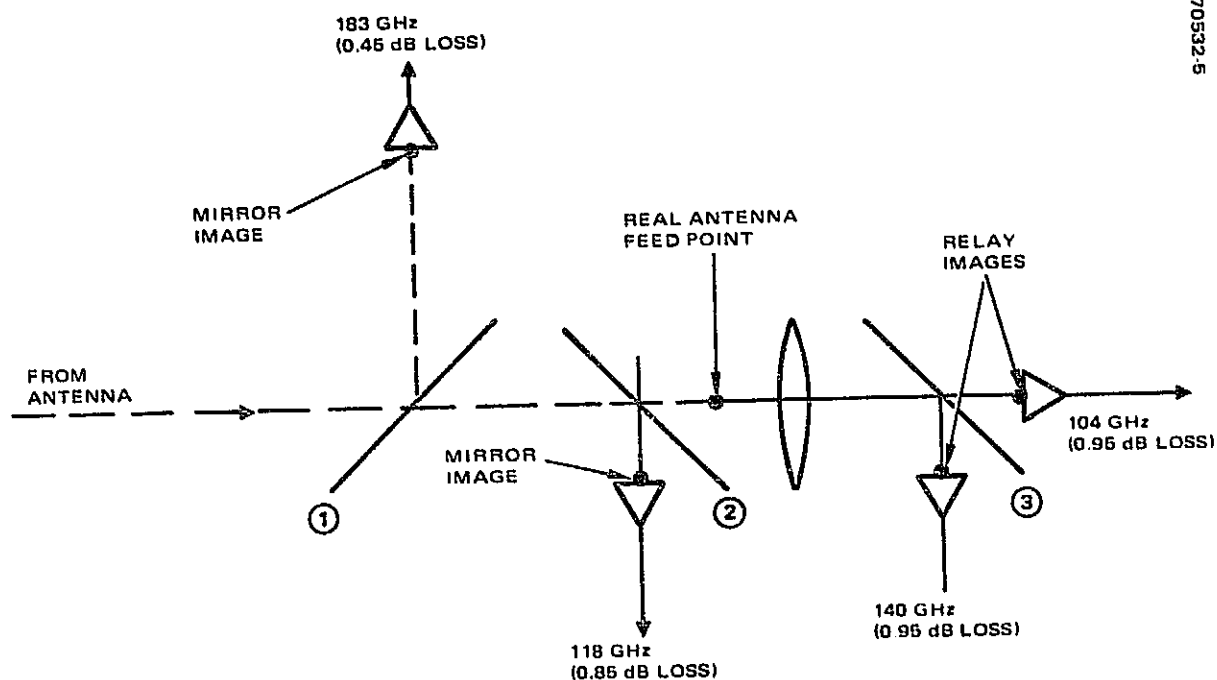


FIGURE 4. QUASI-OPTICAL FEED SCHEMATIC

70532-5

1. Introduction and Summary

1.3 RECOMMENDED RADIOMETER CONFIGURATION

The present study effort has resulted in the selection of an overall configuration which offers excellent theoretical performance and affords the least technical risk.

Millimeter wave technology status offers several approaches to MASR, each having "softness" but all are currently under rapid development. A preferred technical approach is difficult at present, but several configurations can meet MASR requirements. The study effort has resulted in the selection of a tentative configuration which appears to have least technical risk. Technology advances in alternate approaches could change the selection but the present design serves as a model to estimate weight, power, size, and cost of the flight instrument.

The baseline approach utilizes balanced mixers in all bands. IMPATT local oscillators with quasi-optical filters are recommended for 118, 140, and 183 GHz bands, while a TEO oscillator/doubler is recommended for the 104 GHz band. Quasi-optical demultiplexers are recommended in the form of band dropping mirrors for the 183, 140, and 118 GHz bands. These dropping mirrors are either wire grid polarizers, resonant grid dichroic mirrors, or Fabry-Perot filters. An isometric view of the optomechanical assembly is shown in Figure 1. The incoming beam from the antenna reaches a minimum beam waist at the calibration sub-reflector mirror, then focussed sharply to the chopper plane and recollimated by a matched off-axis mirror. The beam is formed and directed through a series of band dropping mirrors and to the balanced mixers for each band. The local oscillator signals are combined with the signal and IF outputs are directed the respective IF networks.

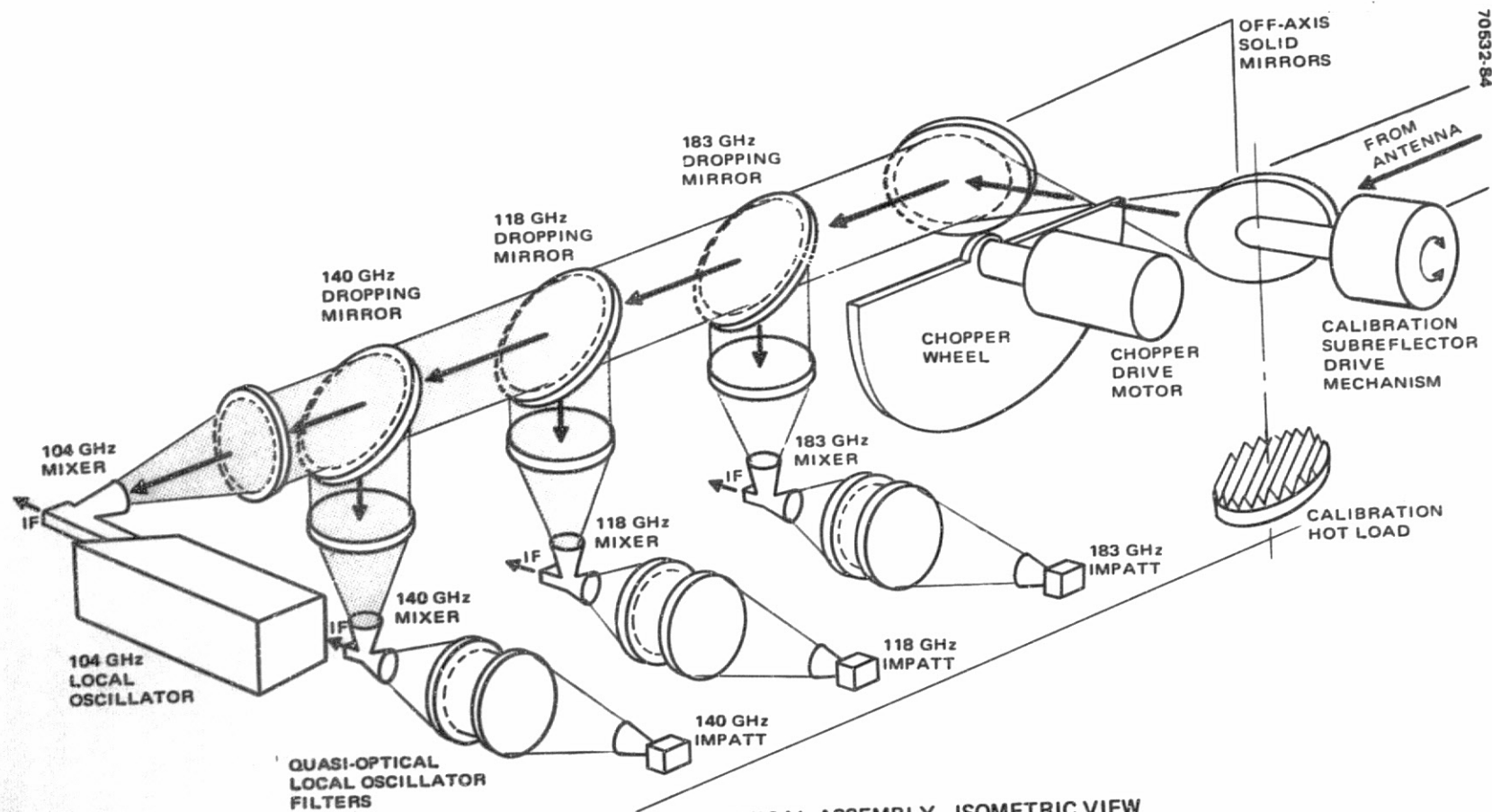


FIGURE 1. MASR OPTOMECHANICAL ASSEMBLY - ISOMETRIC VIEW

2. RECOMMENDED RADIOMETER CONFIGURATION

2.1	PERFORMANCE ESTIMATES	2-2
2.2	RADIOMETER MECHANICAL SUBSYSTEM	2-6
2.3	MASR WEIGHT AND POWER REQUIREMENTS	2-10

2. Recommended Radiometer Configuration

2.1 PERFORMANCE ESTIMATES

Tables 1 through 3 are performance estimates for the four radiometric bands based on state of the art technologies with modest improvements anticipated.

TABLE 1. PERFORMANCE ESTIMATES FOR 118 GHz BAND

Channel Number	f _{IF} , MHz	B _{IF} , MHz	F _{IF} , dB	T _{IF} , °K	ΔT		
					Goal	Dicke (K = 2)	Reference Average (K = 1.4)
T-1	3590	1000	3.3	330	0.2	0.06	0.04
T-2	1930	700	3.3	330	0.2	0.08	0.06
T-3	1140	400	3.5	359	0.2	0.11	0.08
T-4	740	287	3.5	359	0.2	0.14	0.10
T-5	490	212	3.5	359	0.2	0.16	0.11
T-6	315	137	3.0	290	0.2	0.18	0.13
T-7	200	92	3.0	290	0.2	0.22	0.16
T-8	125	57	3.0	290	0.2	0.27	0.19
T-9	80	32	2.5	225	0.2	0.34	0.24
T-10	52.5	22	2.5	225	0.2	0.40	0.29
T-11	30	22	2.5	225	0.2	0.40	0.29
Average					0.2	0.21	0.15
Center frequency					118.75 GHz		
Mixer					Waveguide balanced		
Mixer temperature, T _{MSSB}					1100°K		
Mixer conversion loss, L _m					5 dB		
Local oscillator					IMPATT		
Local oscillator filter					Fabry-Perot parallel plate		
Total IF bandwidth					19 to 4090 MHz		
IF filter					Triplexer/9 way power divider		
Feed loss					1 dB		
Antenna temperature					250°K		
Integration time					1.55 sec		

TABLE 2. PERFORMANCE ESTIMATES FOR 183 GHz BAND

Channel Number	f_{IF} , GHz	B_{IF} , GHz	F_{IF} , dB	T_{IF} , °K	ΔT		
					Goal	Dicke (K = 2)	Reference Average (K = 1.4)
H-1	8.625	2.5	5.5	738	0.2	0.18	0.13
H-2	6.25	2.5	5.1	648	0.2	0.16	0.11
H-3	4.375	1.25	5.3	692	0.2	0.24	0.17
H-4	3.125	1.25	4.9	606	0.2	0.22	0.16
H-5	1.875	1.25	4.3	490	0.2	0.19	0.14
H-6	0.9375	0.625	3.7	389	0.2	0.24	0.17
Average					0.2	0.21	0.15
Center frequency					183.30 GHz		
Mixer					Waveguide balanced		
Mixer temperature, T_{MSSB}					2100°K		
Mixer conversion loss, L_m					8.5 dB		
Local oscillator					IMPATT		
Local oscillator filter					Fabry-Perot parallel plate		
IF bandwidth					0.625 – 10 GHz		
IF filter network					Cascaded duplexers		
Feed loss, L_F					0.5 dB		
Antenna temperature, T_A					250°K		
Integration time					1.0 sec		

70532-97

TABLE 3. PERFORMANCE ESTIMATES FOR 104 AND 140 GHz BANDS

Channel Number	f_{IF} , MHz	B_{IF} , MHz	F_{IF} , dB	T_{IF} , °K	ΔT		
					Goal	Dicke ($K = 2$)	Reference Average ($K = 1.4$)
W-1	500	1000	3.0	290	0.2	0.05	0.04
W-2	500	1000	3.0	290	0.2	0.08	0.06
W-1 center frequency					104 GHz		
W-2 center frequency					140 GHz		
Mixers					Waveguide balanced		
Mixer temperature, T_{MSSB}							
W-1					600°K		
W-2					600°K		
Mixer conversion loss, L_m							
W-1					5 dB		
W-2					6.5 dB		
Local oscillator							
W-1					Gunn doubler		
W-2					IMPATT		
Local oscillator filter							
W-1					Waveguide		
W-2					Fabry-Perot parallel plate		
Total IF bandwidth, each band					1 GHz		
Feed loss, each band					0.95 dB		
Antenna temperature, T_A					250°K		
Integration time							
W-1					1.76 sec		
W-2					1.31 sec		

70532-10T

~~REPRODUCING~~ PAGE BLANK NOT FILMED

2. Recommended Radiometer Configuration

2.2 RADIOMETER MECHANICAL SUBSYSTEM

The mechanical subsystem is designed to be accommodated in the main dish structure of the antenna. The receiver assembly is contained in a package 16 by 16 by 32 inches.

The receiver mechanical assembly includes a structural frame, pivoted mirror for calibration, chopper wheel, hot reference load, demultiplexer, and the four radiometer receivers. Suggested layouts are shown in Figures 1 and 2. The pivoted calibration mirror is shown in position to illuminate the subreflector of the antenna. The mirror is also capable of pointing at the hot reference load and at cold space. The chopper wheel assembly is shown in greater detail in Figure 2. The mirror is an off-axis parabola, bringing the incoming beam from the antenna to a sharp focus at a point where the chopper wheel intersects the beam. The beam waist at focus is 3.6 mm to the 99 percent truncation points. Since the chopper wheel circumference is about 300 mm, the chopper rise and fall transition ratios are given by

$$R_{\text{RISE/FALL}} = \frac{3.6}{150} = 2.4\%$$

Another parabolic segment is matched to the pivoted calibration mirror in such a way as to minimize aberrations. The beam is recollimated and directed to the use of IMPATTS as local oscillators for three bands and a Gunn for the fourth band to illustrate the two available options. Access to the main antenna focus is through a cylindrical space 24 inches in diameter and 36 inches deep (Figures 3 and 4). The surface of the main parabola is provided with a 25 cm diameter circular hole that allows the beam from the antenna to come to focus at a point 16 cm behind the parabolic surface. When located in position, the pivoted subreflector of the receiver lies approximately at the focus of the antenna. A hole will be located in the antenna structure to allow the pivoted calibration subreflector to look at cold space.

The receiver mechanical subsystem occupies only one quadrant of the available space in the rear of the antenna, although the structure extends several inches into the other quadrants to accommodate a symmetrical axial beam.

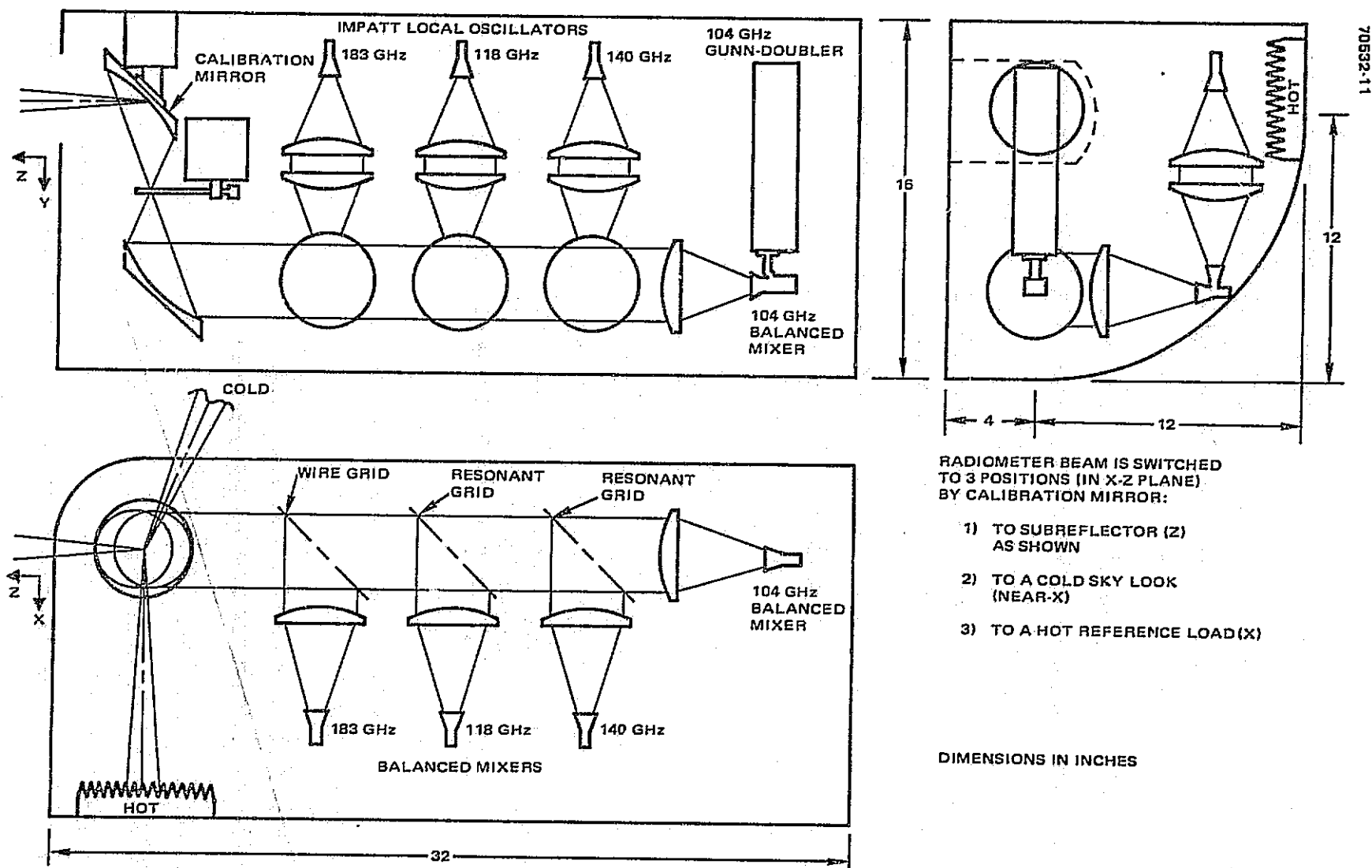


FIGURE 1. RADIOMETER MECHANICAL SUBSYSTEM

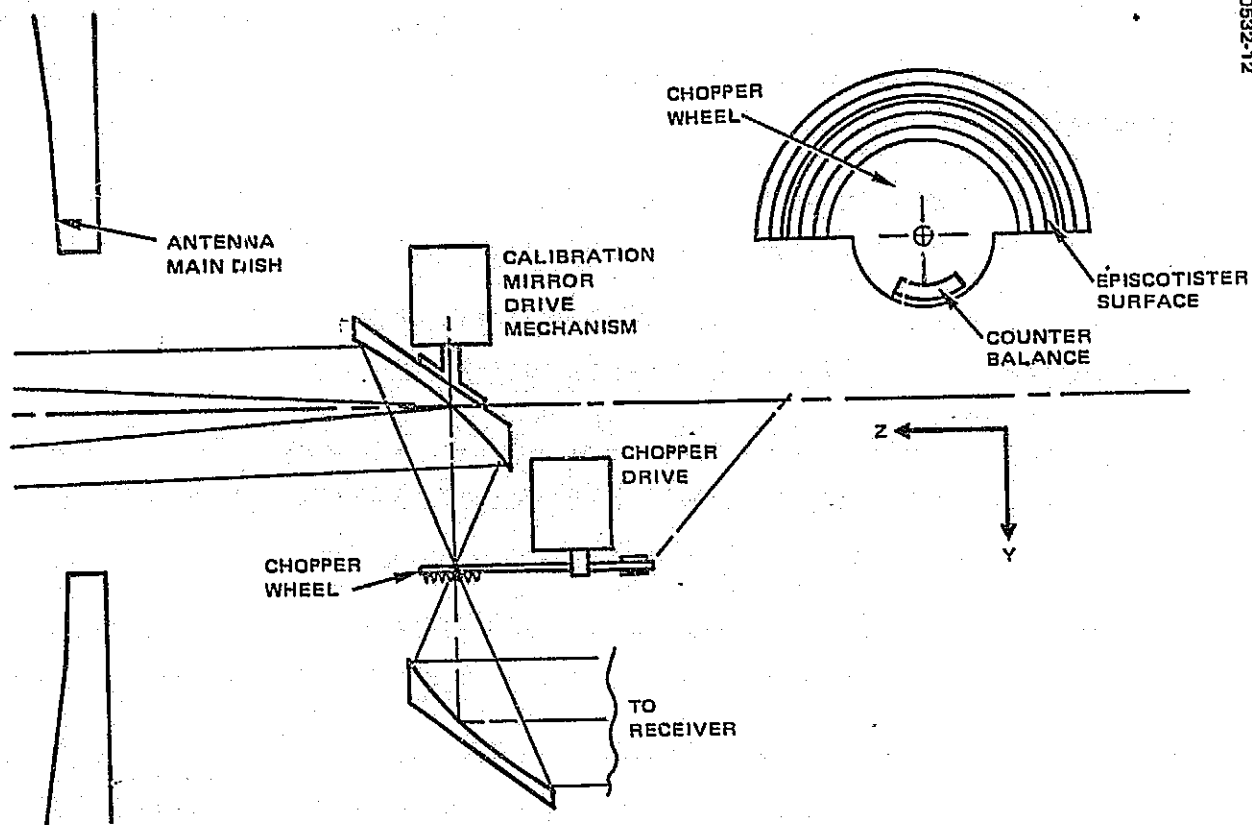


FIGURE 2. CALIBRATION MECHANISM AND CHOPPER WHEEL, TOP VIEW

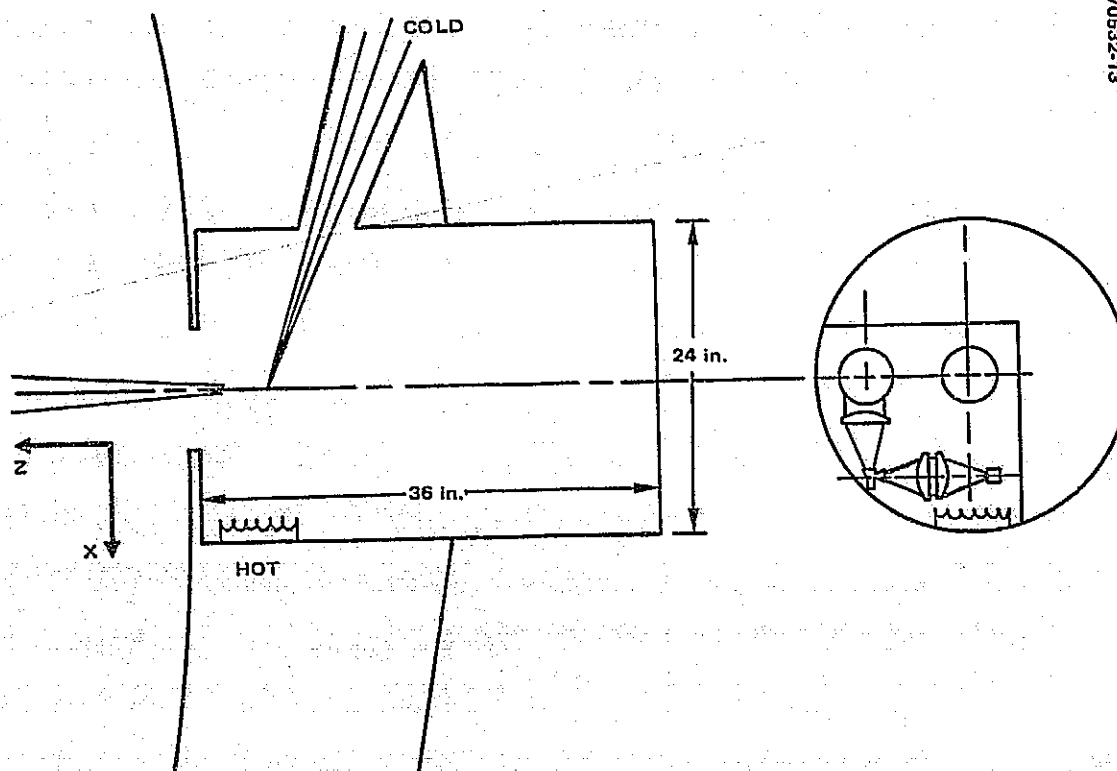


FIGURE 3. APERTURE FOR LOCATION OF RADIOMETER

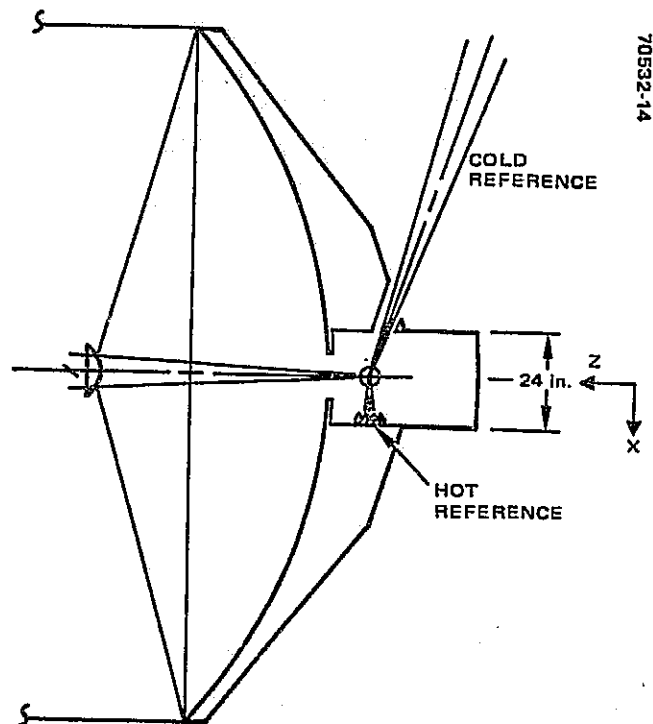


FIGURE 4. ANTENNA SIDE VIEW SHOWING
RADIOMETER CALIBRATION LOADS

2. Recommended Radiometer Configuration

2.3 MASR WEIGHT AND POWER REQUIREMENTS

The weight, power, and thermal/mechanical interface data of the MASR payload are tabulated in Table 1.

TABLE 1. PAYLOAD WEIGHT AND POWER REQUIREMENTS

Weight Tabulation, lb		
Optomechanical subsystem	15.5	
183 GHz receiver subsystem	4.2	
118 GHz receiver subsystem	3.7	
104/140 GHz receiver subsystem	1.8	
Signal processing subsystem	1.0	
Miscellaneous electronics	1.4	
Cables and connectors	4.6	
Chassis	<u>18.0</u>	
Total MASR weight estimate, lb	50.2	
Power Tabulation, W		
Chopper drive	5	
Calibration mirror drive	15	0
Local oscillator sources	45.3	
IF amplifiers	3.2	
Signal processing	13.2	
Power conditioning	16.3	14.3
Total power required, W	98.0 max	81.0 min
Package Dimensions, in.	16 x 14 x 32	
Thermal Load, W	81 to 98	

70532-15T

3. DESIGN ANALYSIS

3.1	MASR CONCEPT	3-2
3.2	NOISE TEMPERATURE	3-4
3.2.1	System Noise Temperature	3-4
3.2.2	Mixer Noise Temperature	3-8
3.2.3	IF Contribution to Noise Temperature	3-10
3.3	RADIOMETER ΔT PERFORMANCE PARAMETRIC TRADEOFFS	3-12
3.4	REFERENCE AVERAGING	3-16
3.5	RADIOMETER CALIBRATION	3-22
3.6	QUASI-OPTICAL DESIGN FORMALISM	3-28
3.7	FABRY-PEROT FILTER DESIGN CONSIDERATIONS	3-36

3. Design Analysis

3.1 MASR CONCEPT

Microwave sounding consists of measuring at selected frequencies the brightness temperature of passive microwave radiation emitted by the earth's atmosphere and surface. Brightness versus frequency near an absorption line yields temperature information as a function of altitude.

The basic equation that relates the brightness temperature of the earth's outgoing radiation to the atmospheric parameters is the microwave radiative transfer equation. (See Appendix A.) The physics of radiative transfer involve: 1) surface radiant emission attenuated by the atmosphere; 2) summation of emissions from all the layers of the atmospheric column (see Figure 1) attenuated by the intervening atmosphere between the layer and outer space; and 3) reflections by the surface of the earth which contribute to the outgoing radiation. In temperature sounding, the brightness versus frequency is measured near the oxygen absorption line at 118 GHz and the temperature profile is inferred from the radiative transfer equation. In humidity sounding, brightness versus frequency is measured near the 183 GHz H₂O line, comparing the resulting temperature profile with that obtained from the oxygen line. Relative humidity as a function of altitude is thereby determined. Figure 1 gives a functional flow diagram for the sounding system.

The radiometer measures directly the antenna temperature given by

$$T_A(f) = \frac{1}{4\pi} \int_{4\pi} G(f, \theta, \phi) T_B(f, \theta, \phi) d\Omega$$

where $G(f, \theta, \phi)$ is the antenna gain function. The calibration of brightness temperature T_B against antenna temperature T_A is done empirically since calculated calibrations based on the antenna gain function are generally inaccurate.

At a given frequency difference from line center, the observed brightness temperature is dominated by a corresponding layer of the atmosphere. For example, the oxygen line center is 118.7 GHz and measurements at this frequency are dominated by gases in the upper atmosphere above 20 km. However, at 118.22 GHz the measurement is dominated by gas at 9 km altitude. The composite weightings for a number of frequencies near the absorption line will thereby yield information about temperature as a function of altitude.

Appendix A presents a more detailed physical analysis of the MASR concept.

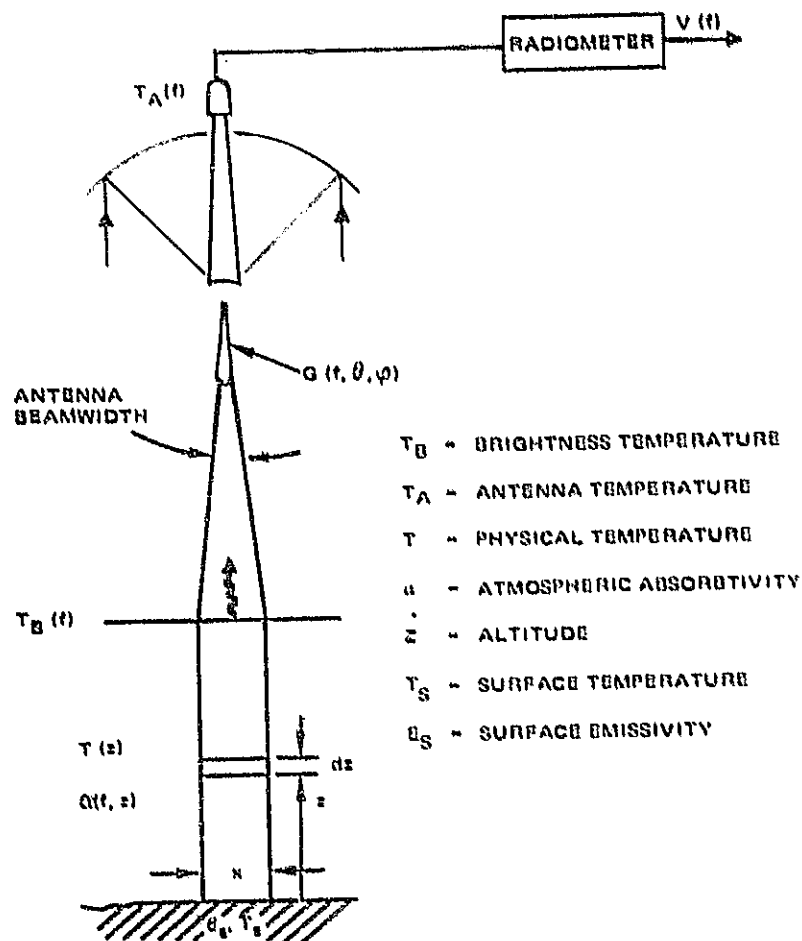
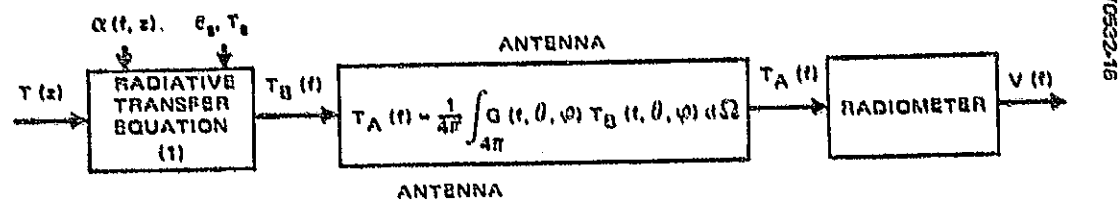


FIGURE 1. MICROWAVE SOUNDING SYSTEM FUNCTIONAL FLOW DIAGRAM

3. Design Analysis

3.2 Noise Temperature

3.2.1 SYSTEM NOISE TEMPERATURE

The sensitivity (temperature resolution) of a radiometer is dependent on the system noise temperature. A discussion of system noise temperature, consistent with convention, is presented.

The receiver noise temperature T_R is the temperature to which the source resistance of an otherwise noiseless receiver must be heated to generate the noise actually produced by a physical and noisy receiver connected to an idealized noiseless source. It therefore represents the noise generated within the receiver referred to its input port.

The system temperature T_S is the sum of T_R and the antenna temperature T_A . T_A results from collected signal energy and is fixed by the average temperature of the earth at about 250°K. T_R is composed of contributions from the mixer and IF, and for a single sideband system

$$T_{RSSB} = T_{MSSB} + L_M T_{IF}$$

T_{MSSB} represents all sources of internally generated mixer noise at signal, image, and IF frequencies referred to the mixer signal input. T_{IF} is the noise temperature of the IF stage, and is multiplied by the mixer signal conversion loss L_M .

In a single sideband receiver system (see Figure 1) where the image sideband is rejected, all sources of internally generated receiver noise are referred to the mixer signal input and the system temperature

$$T_{SSSB} = T_A + T_{RSSB} = T_A + T_{MSSB} + L_M T_{IF}$$

Physically, image rejection can be implemented using a sharp RF filter at the mixer input.

In a double sideband receiver used in the MASR radiometer, signal is received in both the signal and image channels which have similar characteristics ($T_{IM} = T_A$ and $L_{IM} = L_M$) (see Figure 1). The amount of internally generated receiver noise is the same as in the single sideband case, but now its effect is divided between two signal channels with the result that⁽¹⁾

$$T_R = \frac{1}{2} T_{RSSB}$$

$$T_S = T_A + 1/2 (T_{MSSB} + L_M T_{IF})$$

Feed losses between the antenna and the mixer raise the system noise temperature by attenuating the signal and by emitting noise. For equal signal and image channel losses L_F , the system temperature becomes⁽²⁾

$$T_S = T_A = (L_F - 1) T_O + \frac{L_F}{2} (T_{MSSB} + L_M T_{IF})$$

where T_O is the physical temperature of the lossy element. For $T_O = 290^\circ\text{K}$, about 7°K noise is emitted for each 0.1 dB loss. The factor L_F also multiplies the receiver noise temperature, which results in an added contribution for the highest frequency channels of

$$\Delta T_S = \frac{L_F - 1}{2} (T_{MSSB} + L_M T_{IF}) \doteq \frac{1.023 - 1}{2} (4000^\circ\text{K}) = 47^\circ\text{K}$$

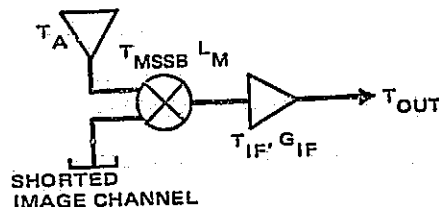
Thus, each 0.1 dB loss between the antenna and mixer adds approximately 54°K to the system noise temperature.

References

1. J.D. Kraus, "Radio Astronomy," McGraw Hill, New York, 1966, pp 265-266.
2. A.J. Giger, et. al., BSTJ, 42, No. 4, P1096, July 1963.

SINGLE CHANNEL

$$T_S = T_A + T_{MSSB} + L_M T_{IF}$$



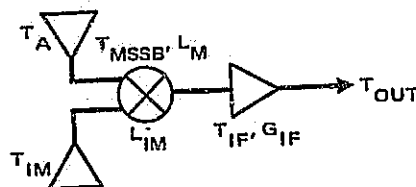
70532-17

DOUBLE-CHANNEL RADIOMETER

$$T_S = T_A + \frac{1}{2} (T_{MSSB} + L_M T_{IF})$$

$$T_{IM} = T_A$$

$$L_{IM} = L_M$$



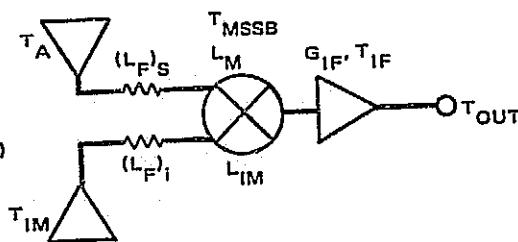
WITH FEED LOSSES

$$T_S = T_A + (L_F - 1) T_O + \frac{L_F}{2} (T_{MSSB} + L_M T_{IF})$$

$$T_{IM} = T_A$$

$$L_{IM} = L_M$$

$$L_{fc} = L_{fs}$$



- $(L_F)_S$ = FEED LOSS FOR SIGNAL CHANNEL
- $(L_F)_I$ = FEED LOSS FOR IMAGE CHANNEL
- T_A = ANTENNA TEMPERATURE FOR SIGNAL CHANNEL
- T_{IM} = ANTENNA TEMPERATURE FOR IMAGE CHANNEL

FIGURE 1. MIXER MODEL EQUIVALENT CIRCUITS

PRECEDING PAGE BLANK NOT FILMED

3.2 Noise Temperature

3.2.2 MIXER NOISE TEMPERATURE

A millimeter wave diode mixer can be characterized by its conversion loss L_M and its equivalent input noise temperature T_{MSSB} . Experiment shows that simple mixer models relating the properties of a diode mixer to those of passive devices have limited usefulness.

Neglecting feed losses, the system temperature of a double sideband receiver was shown to be

$$T_S = T_A + \frac{1}{2} (T_{MSSB} + L_M T_{IF})$$

T_{MSSB} adds directly to system temperature while L_M multiplies the IF noise temperature T_{IF} . In the high frequency MASR channels these two contributions to T_S are of comparable size.

T_{MSSB} can be determined using the relationship for receiver noise temperature

$$T_R = \frac{1}{2} (T_{MSSB} + L_M T_{IF})$$

and measuring T_R , L_M , and T_{IF} . Rapid and accurate techniques for measuring mixer performance have been developed. (1)

An alternate characterization of mixer performance is the mixer output noise temperature ratio t , defined as the ratio of the mixer output noise temperature to its physical temperature when both the signal and image channels are terminated at T_O . (1) When $L_M = L_{IM}$

$$T_{out} = t T_O = \frac{T_O}{L_M} + \frac{T_O}{L_{IM}} + \frac{T_{MSSB}}{L_M}$$

$$T_{MSSB} = (L_M t - 2) T_O$$

A similar concept is that of diode noise temperature T_D , defined as the physical temperature of an attenuator having insertion loss L_M and L_{IM} at the signal and image frequencies and producing the same output temperature as the mixer (2)

$$T_{out} = t T_O = \frac{T_O}{L_M} + \frac{T_O}{L_{IM}} + \left(L_M - 1 - \frac{L_M}{L_{IM}} \right) T_D$$

For $L_M = L_{IM}$

$$T_{MSSB} = (L_M - 2) T_D$$

For fixed values of t or T_D , these mixer characterizations imply similar relationships between T_{MSSB} and L_M , and it is tempting to use these relationships in a tradeoff study. However, data taken with a mixer operating at 87 GHz and at 115 GHz show that these simple models do not adequately describe mixer operation.⁽³⁾ From Figure 1 the data taken at 87 GHz agrees with the mixer noise temperature ratio model, but data at 115 GHz bears no resemblance to the performance predicted by the simple model.

Attempts to relate L_M and T_{MSSB} using a simple model of the millimeter wave mixer are unproductive due to the complex nature of the noise and mixing processes. L_M and T_{MSSB} must both be considered in evaluating mixer performance.

References

1. S. Weinreb and A. R. Kerr, "Cryogenic Cooling of Mixers for Millimeter and Centimeter Wavelengths," IEEE J. Solid State Circuits (special issue on microwave integrated circuits), Volume SC-8, February 1973, pp. 58-63.
2. A. R. Kerr, "Low Noise Room Temperature and Cryogenic Mixers for 80-120 GHz," IEEE Transactions on MTT, Volume MTT-23, No. 10, October 1975, pp. 781-787.
3. D. N. Held, "Analysis of Room Temperature Millimeter Wave Mixers Using GaAs Schottky Barrier Diodes," PhD thesis, Institute for Space Studies, Goddard Space Flight Center X-130-77-6, January 1977.
4. Also see D. N. Held and A. Kerr, IEEE Transactions on MTT, Parts I and II, February 1978.

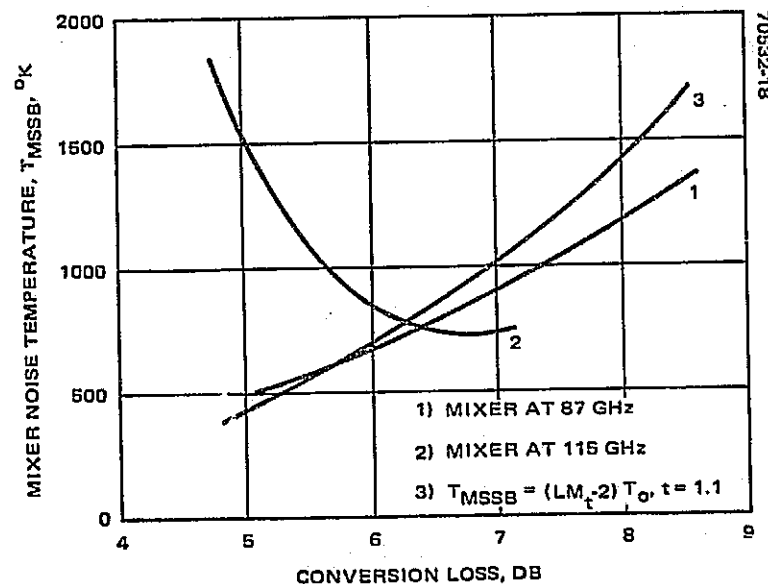


FIGURE 1. EXPERIMENTAL MIXER TEMPERATURE VERSUS CONVERSION LOSS

3.2.3 IF CONTRIBUTION TO NOISE TEMPERATURE

Contributions to the system noise temperature from the IF are determined by the mixer conversion loss as well as the effective noise temperature of the IF network.

The system noise temperature for a radiometer (DSB) was shown to be

$$T_S = T_A + \frac{1}{2} (T_{MSSB} + L_M T_{IF})$$

It can be seen from the above relation that the contribution to the system noise from the IF is dependent on the conversion loss. The receiver noise temperature is

$$T_R = \frac{1}{2} (T_{MSSB} + L_M T_{IF})$$

and

$$T_S = T_A + T_R$$

Now the IF contribution to the receiver noise temperature is simply $L_M T_{IF}$, and is dimensionally equivalent to T_{MSSB} . $L_M T_{IF}$ is plotted in Figure 1 for four values of IF noise temperature as a function of conversion loss. It should be emphasized that the contribution to the system noise temperature for a radiometer application is just half of the plotted values.

The IF noise temperature is generally dominated by the noise figure of the first stage of the IF amplifier. When the gain of the first stage is high and is followed by another stage of gain with reasonably low noise figure, the above rule holds. However, in the proposed radiometer application, the IF may consist of a network of low gain amplifiers and filter banks where the overall effective noise temperature is not entirely determined by the noise figure of the first stage. Figure 2 illustrates one possible IF network where the first stage amplifier is a low gain, broadband amplifier, followed by a demultiplexing filter bank where each band is then amplified separately. Let the noise figure of the first stage be F_A and that of the postamplifiers be F_{BX} , where X is the individual demultiplexed band. A filter loss of L_X is introduced for the X th band. The noise figure for the network is given by

$$F_{AB} = F_A + \left[\frac{L_X F_{BX} - 1}{G_A} \right]$$

and the noise temperature is

$$T_{AB} = (F_{AB} - 1) T_O$$

$$T_{AB} = T_A + \Delta T_E = (F_A - 1) T_O + \left(\frac{L_X F_{BX} - 1}{G_A} \right) T_O$$

where ΔT_E is the excess noise temperature of the network over and above the noise temperature of the first stage. Values of excess noise temperature are plotted as a function of the gain-to-loss ratio of the first stage gain and filter loss for a second stage amplifier noise figure of 2.

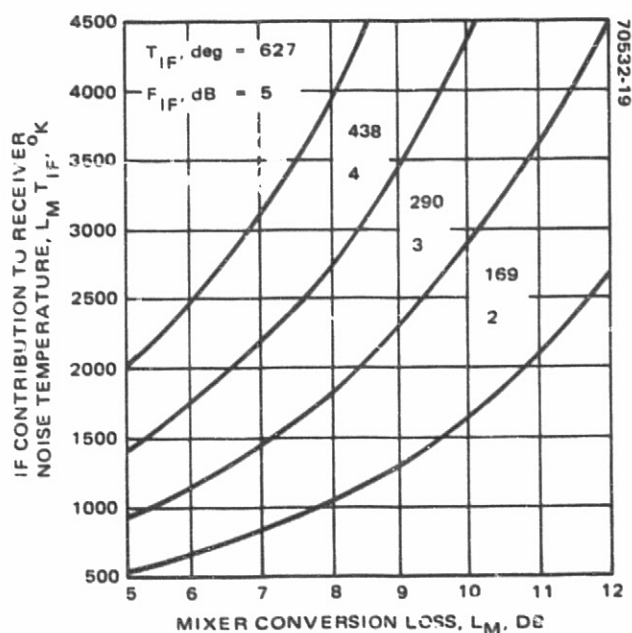


FIGURE 1. IF CONTRIBUTION TO RECEIVER NOISE TEMPERATURE

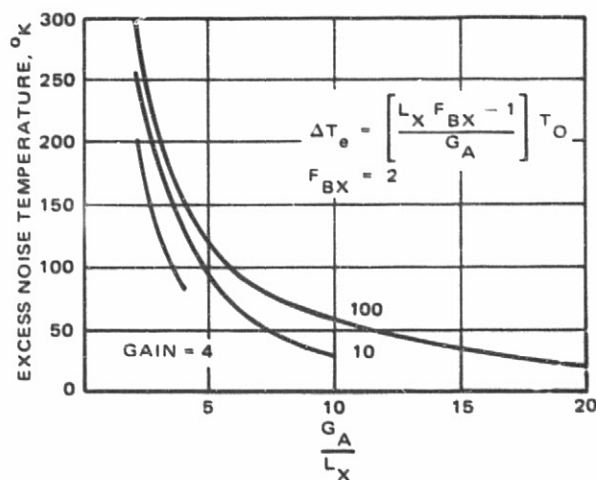
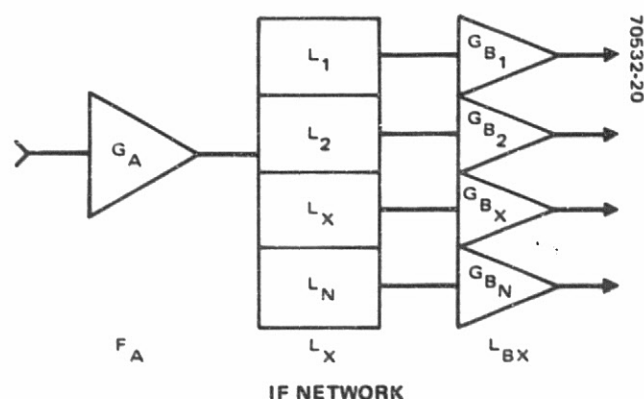


FIGURE 2. EXCESS NOISE TEMPERATURE CONTRIBUTION DUE TO IF NETWORK

3. Design Analysis

3.3 RADIOMETER ΔT PERFORMANCE PARAMETRIC TRADEOFFS

Radiometer ΔT performance is directly proportional to the system noise temperature. Typical values of T_S are estimated for the four radiometric bands and their dependence on feed and mixer conversion losses, and IF noise temperature are analyzed.

The temperature resolution of the Dicke radiometer can be written

$$\Delta T = \frac{2 T_S}{\sqrt{B\tau}}$$

where B is the receiver bandwidth and τ is the radiometer integration time. Since the radiometer performance directly depends on T_S , it is instructive to examine the term closely. System noise temperatures for the four radiometric bands can be calculated from the expression developed in Topic 3.2. Certain assumptions must be made such as antenna noise temperature, mixer conversion loss, IF noise figure, and feed losses. Figure 1 lists tentative calculated values.

The dependence of system noise temperatures and radiometric ΔT upon the included parameters can be analyzed by examining the partial derivatives of T_S with respect to the various parameters. From the general expression for system noise temperature

$$T_S = T_A + (L_F - 1) T_O + L_f T_R$$

$$T_R = \frac{1}{2} (T_{MSSB} + L_M T_{IF})$$

we have

$$T_S = T_A + (L_F - 1) T_O + \frac{L_F}{2} [T_{MSSB} + L_M T_{IF}]$$

The relative magnitudes of the partial derivatives tells which parameters are most important in achieving and maintaining a low noise temperature. Taking a particular example of the 183 GHz band and the parameter values given in Figure 1, the partial derivatives and their numerical values are

$$\frac{\partial T_S}{\partial L_F} = T_O + \frac{T_R}{2} = 3818^\circ$$

$$\frac{\partial T_S}{\partial F_{IF}} = \frac{L_F}{2} L_M T_O = 1152^\circ$$

$$\frac{\partial T_S}{\partial L_M} = \frac{L_F}{2} T_{IF} = 393^\circ$$

$$\frac{\partial T_S}{\partial F_{MSSB}} = \frac{L_F}{2} T_O = 163^\circ$$

In two of the expressions, noise figures were used instead of noise temperatures to make all of the expressions dimensionally equivalent. The results indicate that feed losses are of greatest importance, followed by IF noise figure, mixer conversion loss, and mixer noise figure.

It should be emphasized that these expressions are for small incremental effects. Large changes in parameters, such as 3 dB increments, give substantially different results. Figure 1 illustrates how the ΔT values for the 183 GHz band change for a large (3 dB) change in mixer conversion loss, IF noise figure, and feed losses. In the left column of Figure 1, the ΔT goals are listed for the six H₂O channels; in the right columns the calculated values of ΔT are listed for each of three values of mixer conversion loss. The value of $L_M = 8.5$ dB is selected as typical. In the next group of computations, the effects of increasing the IF noise figure are listed. In the final grouping of computations, the effects of increased feed losses are shown.

Conclusions are that in striving to achieve ultimate performance, feed losses have the greatest incremental effect. However, gross changes in any of the parameters have almost a direct effect on performance.

70533-21

H ₂ O CHANNEL	ΔT GOAL	F _{IF}	CONDITIONS			ΔT FOR		
			T _{IF}	B _{IF}	L _F	L _M = 7	L _M = 8.5	L _M = 10
H-1	0.2	5.5	739	2.5	0.5	0.10	0.12	0.16
H-2	0.2	5.1	648	2.5	0.5	0.09	0.11	0.14
H-3	0.2	5.3	693	1.25	0.5	0.13	0.17	0.21
H-4	0.2	4.9	606	1.25	0.5	0.13	0.16	0.19
H-5	0.2	4.3	491	1.25	0.5	0.11	0.14	0.17
H-6	0.2	3.7	390	0.625	0.5	0.14	0.17	0.20
 								
H-1		8.5	1763				0.24	
H-2		8.1	1582				0.22	
H-3		8.3	1671				0.32	
H-4		7.9	1498				0.29	
H-5		7.3	1267				0.26	
H-6		6.7	1066				0.32	
								
H-1		5.5			3.5		0.25	
H-2		5.1			3.5		0.23	
H-3		5.3			3.5		0.33	
H-4		4.9			3.5		0.31	
H-5		4.3			3.5		0.27	
H-6		3.7			3.5		0.34	

ΔT = RADIOMETER SENSITIVITY
 F_{IF} = IF NOISE FIGURE, dB
 T_{IF} = IF NOISE TEMPERATURE, °K
 L_F = FEED LOSS, dB
 L_M = MIXER CONVERSION LOSS, dB
 B_{IF} = IF BANDWIDTH, GHz
 T_A = 250°K
 T_{MSSB} = 2100°K
 K = 1.4 (REFERENCE AVERAGING)

$$\Delta T = \frac{KT_s}{\sqrt{B_{IF}}}$$

$$T_s = T_A + (L_F - 1) T_O + \frac{L_F}{2} (T_{MSSB} + L_M T_{IF})$$

FIGURE 1. RADIOMETER ΔT TRADES AT 183 GHz BAND

3. Design Analysis

3.4 REFERENCE AVERAGING

Reference averaging can reduce the Dicke radiometer K factor from 2 to some value of the order 1.3 to 1.6 depending on chopping fraction and gain stability. A summary of the detailed analysis given in Appendix B appears below.

The temperature resolution of a radiometer is determined by errors from two sources: 1) random fluctuations inherent in the square law detection of noise power in a frequency band and 2) temporal variations in gain. The use of "reference averaging" to improve a Dicke radiometer, first proposed by Bremer and Prince, (1) has the effect of reducing only the first type of errors. The technique reduces the variance in the reference voltage by increasing its integration time for postdetection smoothing, which leads to a decrease in the overall variance of the differential output voltage. In a switching radiometer with 50 percent duty cycle instead of integrating over one-half the dwell time (τ_p) for a resolution element or "pixel," we can integrate over $N\tau_p/2$ by averaging the reference samples in N consecutive pixels. Reference averaging can be implemented using a digital signal processor to store the required signal and reference data, and to perform the necessary averaging numerically. Such an extension of integration time is possible for the reference samples because the reference temperature is essentially constant, whereas for the antenna samples the integration time must be limited to $\tau_p/2$ to resolve the variation of scene temperature from pixel to pixel. For this case, the antenna temperature error δT_{An} due solely to detector fluctuations is given in the Bremer and Prince report. (2)

$$\delta T_{An} = \frac{T_S}{\sqrt{B\tau_p}} \cdot K = \frac{T_S}{\sqrt{B\tau_p}} \sqrt{2 + \frac{2}{N}} \quad (1)$$

where T_S is the system noise temperature, B is the IF or predetection bandwidth, τ_p is the pixel dwell time, N is the number of pixels over which the reference samples are averaged, and K is the radiometer sensitivity factor. In Equation 1, square wave modulation and detection are assumed.

To determine how large N should be and to estimate the possible improvement, a detailed study was made to understand how reference averaging affects gain variability errors (see Appendix B). We derived the following expression for the antenna temperature error δT_{Ag} due to gain variation in a reference averaging radiometer for symmetrical chopping and averaging

$$\delta T_{Ag} = \frac{1}{\bar{G} + \frac{\gamma_1}{2}} \left[\left(\bar{G} + \frac{\delta \gamma_1}{2} \right) (T_A - T_D) + \delta \gamma_N (T_D + T_R) \right] \quad (2)$$

where $\bar{G}$ is the average of the time varying gain $G(t)$ within a pixel dwell time; T_A , T_D , and T_R are the antenna, Dicke reference, and receiver noise temperatures, respectively. γ_N is an integral involving $G(t)$ defined as

$$\gamma_N = \frac{2}{\tau_p} \int_{N\tau_p} C_N(t) G(t) dt \quad (3)$$

γ_N depends on the parameters N (N is the number of pixels being averaged and is a positive integer ≥ 1), τ_c , and τ_p via the square wave function $C_N(t)$, which is a reference waveform synchronized with the temperature modulation that coherently multiplies the detector output in our model (Figure 1) of a reference averaging system. $C_N(t)$ may be described as a sequence of rectangular pulses of width $\tau_c/2$ and heights equal to $+1$, $-1/N$, or zero, symmetrically distributed over a time interval of $N\tau_p$; outside this interval $C_N(t) = 0$. The $C_N(t)$ functions for $N = 1, 3$, and some arbitrary integer > 3 are illustrated in Figures 2b, c, and d, respectively, for the case $\tau_c = 0.5 \tau_p$. It is assumed that the N consecutive pixels of reference samples being averaged are time symmetrically distributed about the center of the observed pixel interval. In Equation 2, $\delta\bar{G}$, $\delta\gamma_1$, and $\delta\gamma_N$ are the variations in the gain dependent quantities $\bar{G}$, γ_1 , and γ_N from their values in the last calibration.

Typically, $|\gamma_1| \ll |\bar{G}|$, so the first term in Equation 2 is essentially $(\delta\bar{G}/\bar{G})(T_A - T_R)$, which is the well known formula for the gain variation error in a Dicke radiometer. The second term $(\delta\gamma_N/\bar{G})(T_D + T_R)$, which we denote by $\delta\hat{T}_A$, is negligible in a conventional Dicke where $N = 1$, but grows rapidly with increasing N when reference averaging is used. Assuming a specific model for $G(t)$, the error $\delta\hat{T}_A$ is evaluated in Appendix B for typical system parameters, and is plotted as a function of N in Figure 3 for different percent gain fluctuations at 0.01 Hz. The curves in Figure 3 assume a system noise temperature of 3500 °K, $\tau_p = 1$ sec, and $\tau_c = 0.1 \tau_p$.

In a reference averaging radiometer, the gain variation error is not zero even when $T_A = T_D$, since only the first term in Equation 2 vanishes but not the second. This is because some reference samples are taken at a time up to $1/2 N\tau_p$ from the T_A samples. This case corresponds to minimum gain variation error

$$(\delta T_{Ag})_{\min} = \frac{1}{\bar{G} + \frac{\gamma_1}{2}} \delta\gamma_N (T_D + T_R) \\ = \delta\hat{T}_A$$

and is given by Figure 3. Since δT_{Ag} increases with N while δT_{An} decreases with N , there is some optimal value of N which minimizes the combined error $\delta T_A = \sqrt{(\delta T_{An})^2 + (\delta T_{Ag})^2}$, where the two types of error are assumed to be uncorrelated. Figure 4 (derived in Appendix B) shows the generalized K factor K' as a function of N , where K' is defined by setting $\delta T_A = T_{SN}/\sqrt{B\tau_p} K'$ with δT_{An} and δT_{Ag} being derived from Equation 1 and Figure 3, respectively.

Further reduction in detector noise is possible by incorporating asymmetric chopping with reference averaging. If the fraction of time for antenna sampling is lengthened from $0.5\tau_p$ to $q\tau_p$. The reference integration time per pixel is shortened from $0.5\tau_p$ to $(1-q)\tau_p$, but this is counteracted by averaging over N pixels to give a total reference integration time of $N(1-q)\tau_p$. Instead of Equation 1, the error δT_{An} is now a function of N and q

$$\delta T_{An} = \frac{T_{SN}}{\sqrt{B\tau_p}} K(N, q) = \frac{T_{SN}}{\sqrt{B\tau_p}} \sqrt{\frac{1}{q} + \frac{1}{N(1-q)}} \quad (4)$$

70532-22

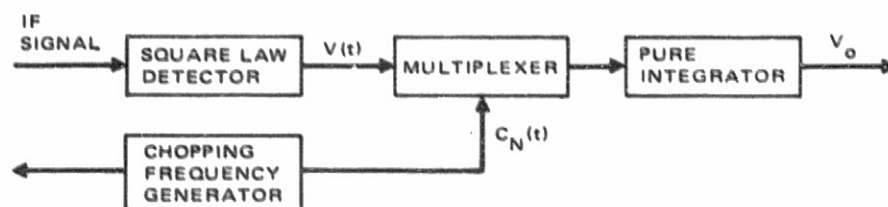


FIGURE 1. POSTDETECTION BLOCK DIAGRAM OF REFERENCE AVERAGING RADIOMETER

70532-23

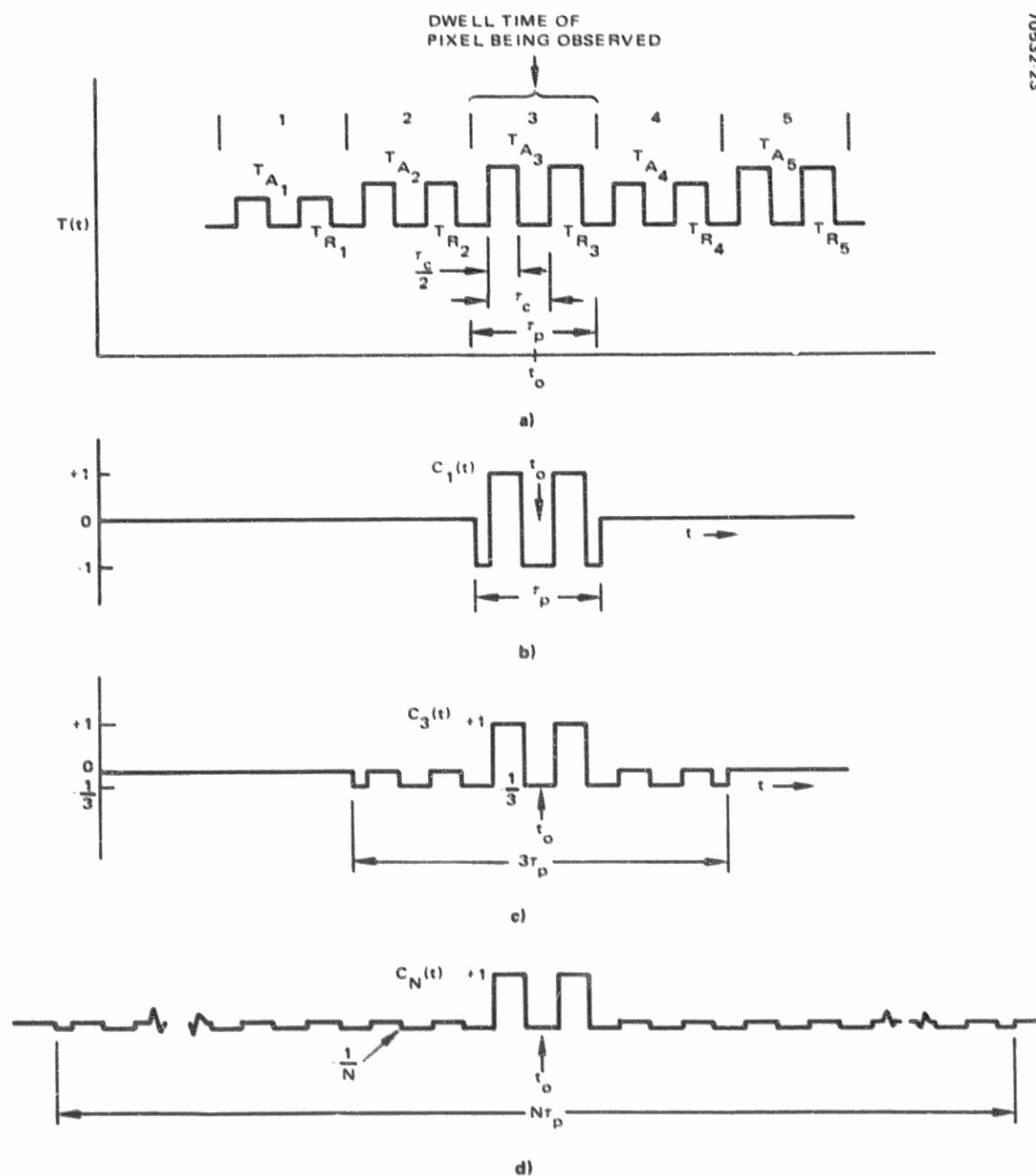


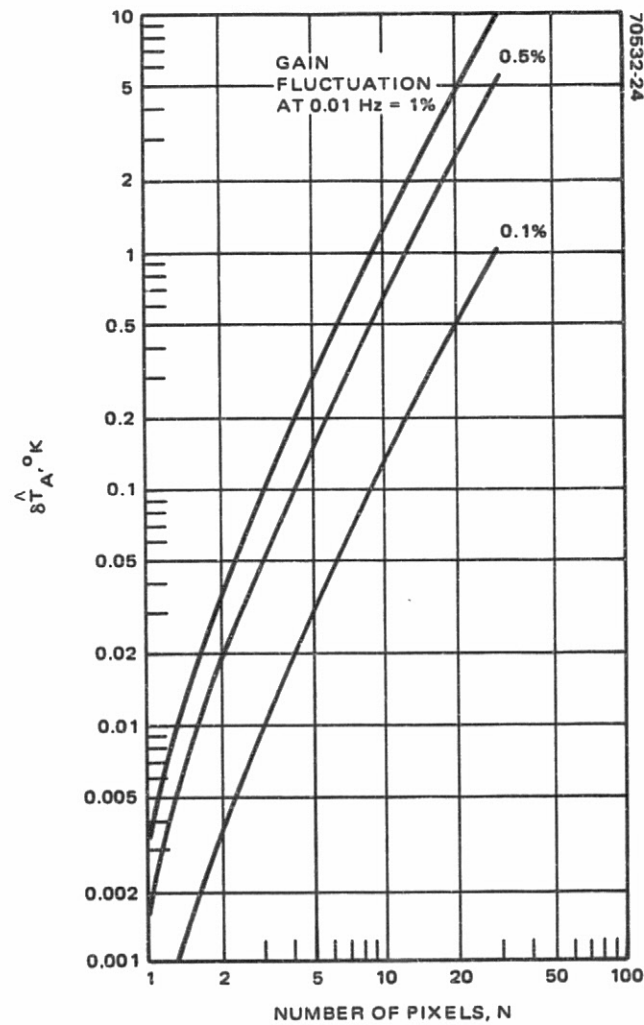
FIGURE 2. RADIOMETER OUTPUT WAVEFORMS

To do a similar analysis for this case, we need to know also δT_{Ag} as a function of N and q . This will be derived in a later study. Knowing $\delta T_{Ag}(N, q)$ and Equation 4, we can then determine values of N and q that minimize the combined error δT_A to find the maximum improvement achievable by reference averaging with asymmetric chopping. For any fixed q , we expect $K'(N)$ for asymmetric chopping to behave in a manner similar to Figure 4.

On the basis of Figures 3 and 4, we tentatively estimate very roughly that the K factor can be reduced by reference averaging to some value of the order of 1.3 to 1.6, depending on the chopping fraction used and the gain stability achievable. The upper value is based on symmetric chopping and the data in Figure 4. The lower value is based on asymmetric chopping and the assumption that gain variation errors can be kept small up to $N = 10$.

References

1. J.C. Bremer, R.E. Prince, J.C. King, and C.D. Calhoun, IEEE Antennas and Propagation Symposium Session ISSO, Stanford, California, 21 June 1977, p. 505.
2. J. Bremer and R.E. Prince, "Microwave Radiometer Design Study," Operations Research, Inc. Report to GSFC, contract NAS 5-23437, Mod 11, January 1977.



PRECEDING PAGE BLANK NOT FILMED

FIGURE 3. ERROR δT_A VERSUS N

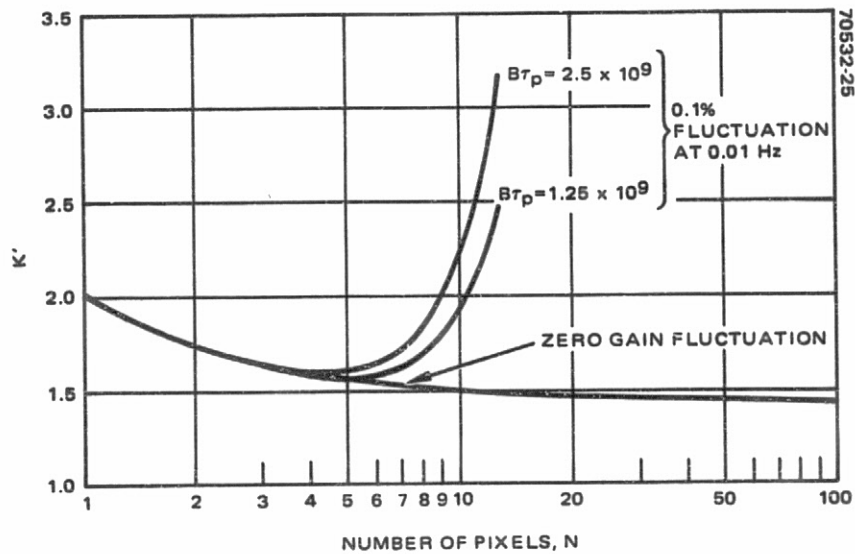


FIGURE 4. GENERALIZED K FACTOR (COMBINED ERROR FROM DETECTOR NOISE AND GAIN FLUCTUATIONS) VERSUS N ($T_D = T_A$)

3.5 RADIOMETER CALIBRATION

Radiometer calibration accuracy, exclusive of the antenna, is better than $\pm 1^\circ\text{K}$ in all channels.

The accuracy with which remote temperature can be radiometrically measured is dependent on radiometer calibration accuracy, radiometer linearity, and size and rate of variations in radiometer gain, noise level, and bandwidth during the time interval between calibration and data collection. The calibration accuracy is examined below.

The dc component of the radiometric output may be represented by

$$V_x = a + bT_x + cT_x^2 + \dots$$

where a and b represent the dc bias and gain constants of the radiometric receiver and T is the input temperature to the radiometer. Nonlinear terms cT_x^2 , etc., should be negligible. The radiometer will be carefully tested for linearity before launch and any required corrections will be determined.⁽¹⁾

Let T_a represent the antenna temperature when the antenna is viewing the earth, T_1 the antenna temperature when the antenna is viewing cold space, and T_2 the temperature of the hot reference load. Given the output voltage for these three situations, the antenna temperature is

$$T_a = \left(\frac{V_2 - V_a}{V_2 - V_1} \right) T_1 + \left(\frac{V_a - V_1}{V_2 - V_1} \right) T_2$$

From this and the linear relationship between voltage output and input temperature, the sensitivity of T_a to errors in T_2 and T_1 may be written

$$\Delta T_a = \left(\frac{T_2 - T_a}{T_2 - T_1} \right) \Delta T_1 - \left(\frac{T_1 - T_a}{T_2 - T_1} \right) \Delta T_2$$

Examination of this equation shows that the sensitivity to errors in T_1 is proportional to $T_2 - T_a$, and the term vanishes when $T_2 = T_a$. Errors due to uncertainty in measuring the radiometer output voltage because of analog-to-digital quantization and nonlinearity will be smaller than the radiometer ΔT and can be neglected.

Radiation at cold space and hot reference temperatures is reflected into the feed using a movable calibration mirror (Figures 1 and 2). The temperature of cold space is known and the efficiency of the cold look calibration will be measured before launch.

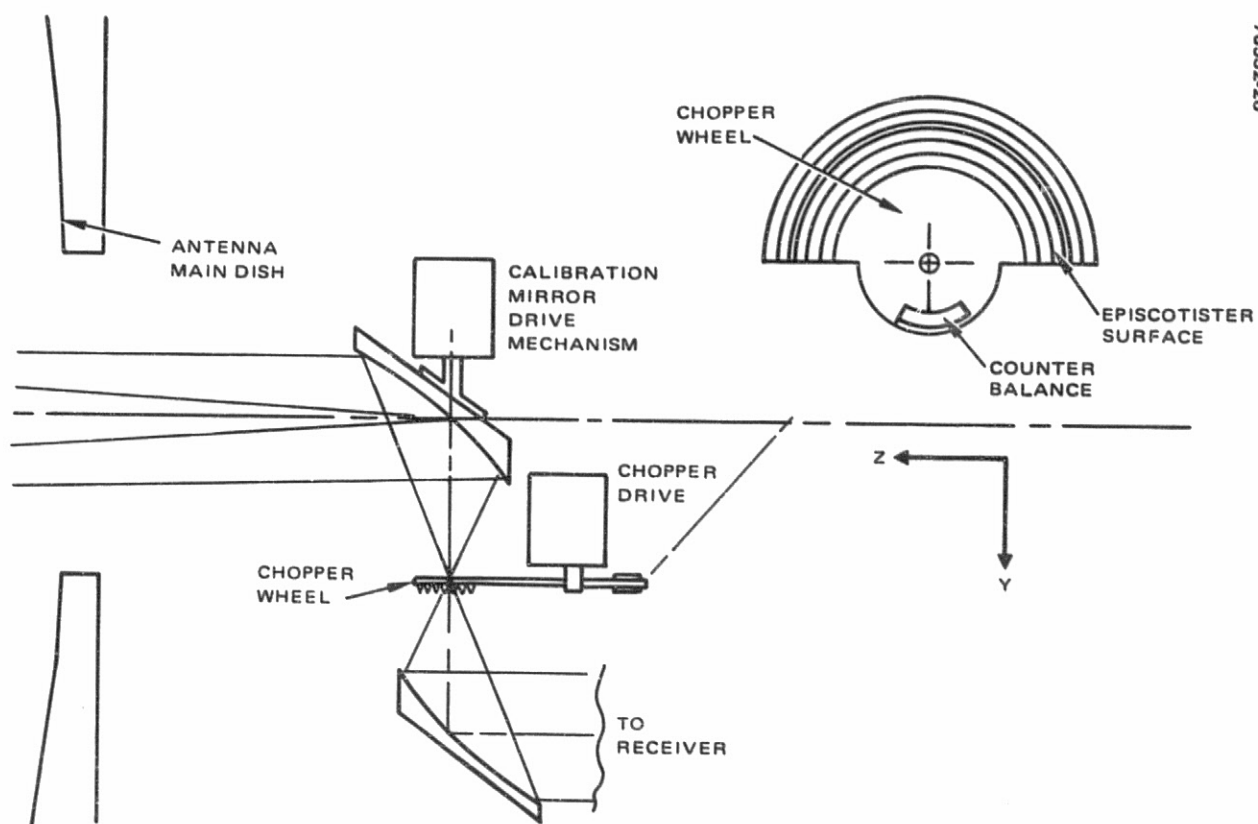
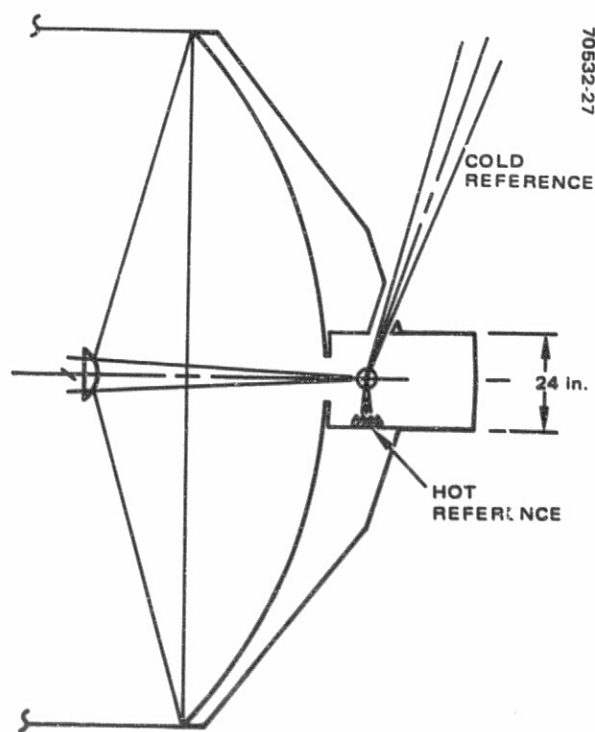


FIGURE 1. CALIBRATION MECHANISM AND CHOPPER WHEEL, TOP VIEW

FIGURE 2. ANTENNA SIDE VIEW SHOWING
RADIOMETER CALIBRATION LOADS

At frequencies above 100 GHz, the radiometric temperature scale undergoes a constant depression in relation to the physical temperature scale, and at temperatures below about 10°K is nonlinear (see Figure 3). This behavior is due to inaccuracies in the Rayleigh-Jeans approximation to Planck's radiation law

$$P = KT \int_B \left(\frac{hf/kT}{\exp(hf/kT) - 1} \right) df \doteq KTB$$

where

P = radiation power emitted by matched load

k = Boltzman's constant

h = Planck's constant

T = temperature

f = frequency

B = bandwidth of interest

Since we calibrate our system using a thermal emitter, we are calibrating to the physical temperature scale for temperatures above 10°K. However, below 10°K the concept of a linear temperature scale is no longer valid, and at 183 GHz, the zero radiation corresponds to a "temperature" of 4.1°K on our scale. Therefore, for calibration purposes we assign a temperature T_1 to cold space of 4.2°K when it is radiating at a Rayleigh-Jeans radiometric temperature of 0.1°K.

The radiometer calibration procedure requires the use of a high emissivity microwave absorber at an accurately known temperature to provide a hot reference. To achieve the required performance, Brewster angle impedance matching techniques developed by Richard Iwasaki at the Jet Propulsion Laboratory are used.⁽²⁾

The hot target is of a ridge design, using an iron filled, high dielectric constant epoxy absorber is cast directly onto an aluminum baseplate (Figure 4). Baseplate fins protruding into the absorbing material are used to decrease temperature gradients from the tip to the root of the ridges. Platinum resistance sensors accurate to $\pm 0.1^\circ\text{K}$ are used to monitor target temperature.

The Brewster angle for the high dielectric constant materials used is relatively independent of frequency, allowing the absorber to be used for calibration over wide frequency ranges.

The measured emissivity of the calibration target built for use in Nimbus-F by JPL was 99.6 percent.

A Dicke radiometer using a chopper wheel reference load is used to reduce the effect of small variations in radiometer performance between calibrations.⁽³⁾

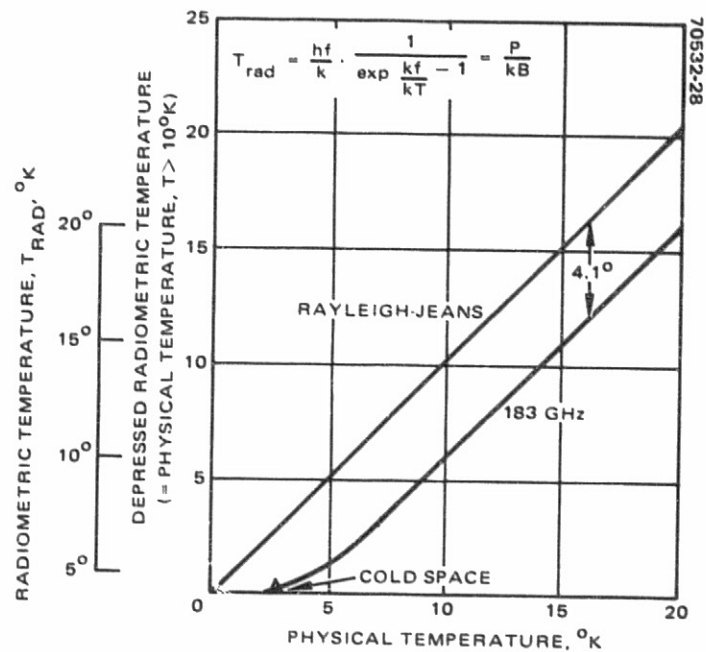


FIGURE 3. RELATIONSHIP BETWEEN PHYSICAL AND RADIOMETRIC TEMPERATURES AT 183 GHz

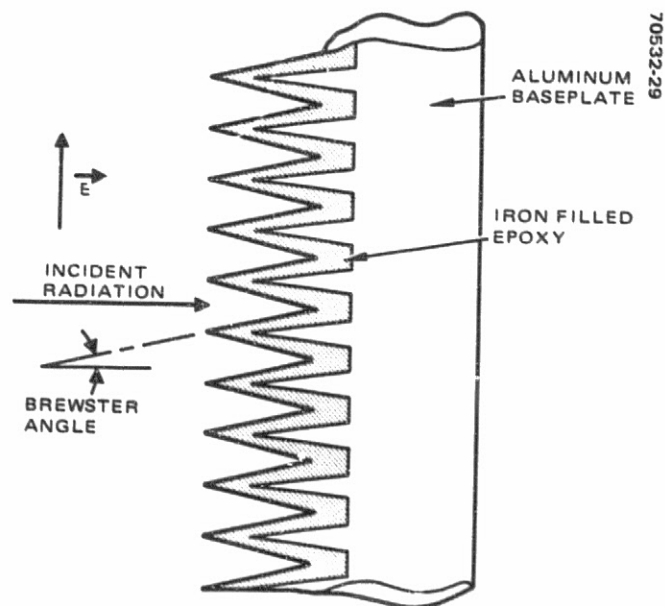


FIGURE 4. HOT REFERENCE, SIDE VIEW

To calculate an estimate of MASR radiometer temperature accuracy, the root sum of the squares of errors in T_A due to estimated uncertainty in the calibration temperatures and due to the radiometer ΔT will be taken.

The temperature of cold space is known, but uncertainty in the efficiency of the cold look aperture will introduce some error

$$T_1 = 4.0 \pm 2^\circ\text{K}$$

$$\text{Radiometer } \Delta T = 0.2^\circ\text{K}$$

$$\Delta T_A^2 = \left[\frac{T_2 - T_A}{T_2 - T_1} \Delta T_1 \right]^2 + (\Delta T)^2$$

PRECEDING PAGE BLANK NOT FILMED

$$\Delta T_A = \left[(0.34)^2 + (0.2)^2 \right]^{1/2} = 0.39^\circ\text{K}$$

The emissivity of the hot absorber will be >99 percent, the temperature will be monitored to $\pm 0.1^\circ\text{K}$, and temperature gradients will be $< 0.2^\circ\text{K}$.

$$T_2 = \sim 300 \pm 0.5^\circ\text{K}$$

$$\Delta T_A = 0.47^\circ\text{K}$$

The total uncertainty in T_A includes contributions from errors in the two calibration temperatures and a contribution from the radiometer ΔT

$$(\Delta T_A)^2 = (0.39^\circ\text{K})^2 + (0.47^\circ\text{K})^2 + (0.2^\circ\text{K})^2$$

$$\Delta T_A = 0.64$$

Radiometer calibration accuracy better than 1°K can be expected in all channels.

References

1. Tiros-N Microwave Sounding Unit Pre-Ship Review, Goddard Space Flight Center, Greenbelt, Maryland, 16 August 1977.
2. Iwasaki, Richard, "Microwave Blackbody Radiometer Calibration Standard," Microwave Workshop, NELC, November 1976.
3. Kraus, J.D., "Radio Astronomy," McGraw-Hill, New York, 1966.

3. Design Analysis

3.6 QUASI-OPTICAL DESIGN FORMALISM

Design formalism using ray matrices and Gaussian beam illumination profiles provide an extremely versatile design approach to quasi-optical systems. Unconditional stability of the system assures negligible loss due to diffraction and radiation from quasi-optical open structures.

Quasi-optical structures are preferred over waveguide structures for diplexer filters, local oscillator filters, and general signal management because of inherent higher Qs and lower losses of the structures. Conventional optical design techniques are applicable which may include extreme geometries and arbitrary illumination. However, with some qualifications (paraxial rays and Gaussian illumination) great simplification of design can be achieved. The methods described were developed by H. Kogelnik and T. Li as a means of simplifying the description and design of laser resonators.* A summary of the Kogelnik and Li paper is given below.

Ray Matrices. The passage of paraxial rays through linear optical structures can be described by ray transfer matrices. Paraxial rays are rays that have small slopes with respect to the optic axis. For a given position along the z-axis (propagation axis), the ray is characterized by its distance r from the optic (z)-axis, and by its angle or slope r' with respect to that axis (see Figure 1). The ray matrix, or ABCD matrix, relates the input quantities r_1 and r_1' to the output quantities r_2 and r_2' by

$$\begin{bmatrix} r_2' \\ r_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} r_1' \\ r_1 \end{bmatrix}$$

Each element of a ray matrix has physical significance

A = lateral magnification

B = equivalent length

C = equivalent focal length, $-\frac{1}{f}$

D = angular magnification.

The principal planes are given by

$$h_1 = (D-1)/C$$

$$h_2 = (A-1)/C$$

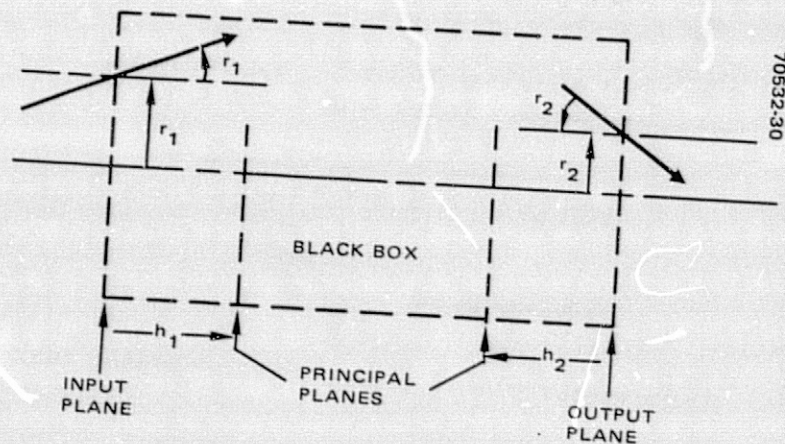


FIGURE 1. RAY MATRIX DIAGRAM

* See Appendix C, Herwig Kogelnik and Tingye Li, Applied Optics, 5, p 1550, October 1966.

Ray matrices are unimodular, that is, their determinants equal unity. Thus, a ray matrix M_s of complex optical structures is built up by cascading matrices of individual elements M_i .

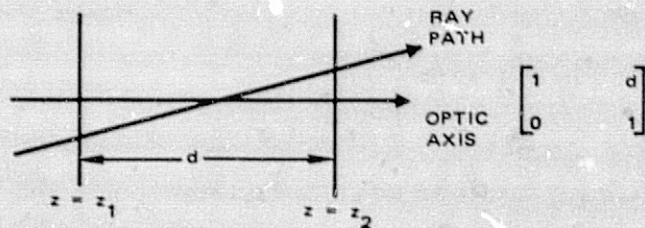
$$M_i = \begin{bmatrix} A & B \\ C & D \end{bmatrix}_i \quad \text{DET } M_i = 1$$

$$\text{DET } M_i = 1$$

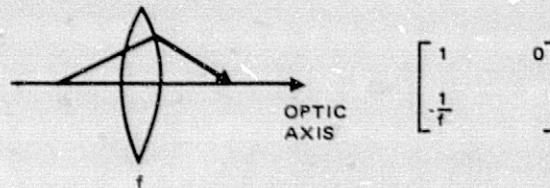
$$M_s = M_1 M_2 M_3 \dots M_i$$

Ray matrices for four common optical elements are given in Figure 2.

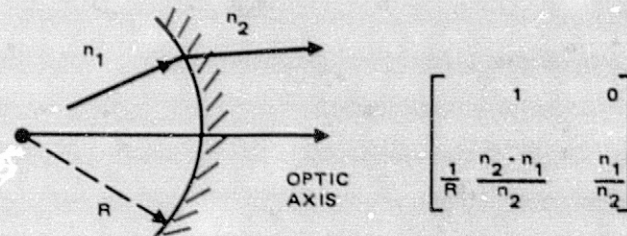
STRAIGHT SECTION
OF LENGTH d



THIN LENS - FOCAL
LENGTH f
• $f > 0$, CONVERGING LENS
• $f < 0$, DIVERGING LENS



SPHERICAL DIELECTRIC
INTERFACE OF RADIUS R
• POSITIVE IF RAY ENCOUNTERS
CONCAVE INTERFACE OR $R = \infty$
• R NEGATIVE IF SURFACE IS
CONVEX



SPHERICAL MIRROR - RADIUS
OF CURVATURE R
• SIGN OF R SAME AS ABOVE

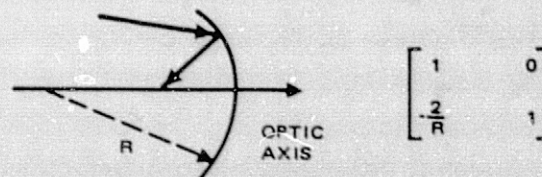


FIGURE 2. RAY MATRICES FOR COMMON OPTICAL ELEMENTS

Gaussian Beams. Gaussian beams satisfy the scalar wave equation and, in many respects, behave similarly to uniform plane waves. However, their intensity distributions are not uniform but are concentrated near the axis of propagation, and their phase fronts are slightly curved. Gaussian beams are mathematically tractable and physically reasonable. For example, fields emitted by lasers are Gaussian; TEM waves in cylindrically symmetric waveguides and HE_{11} modes in any waveguide are very nearly Gaussian. Fields produced by scalar horns can be tailored to approximate Gaussian illumination patterns. A typical Gaussian illumination profile is illustrated in Figure 3.

1) Beam Radius. The width of the beam is measured by the "beam radius" w , defined as the distance at which the field amplitude is $1/e$ times that on axis.

2) Beam Expansion. A Gaussian beam contracts to a minimum diameter $2w_0$ at the beam waist where the phase front is a plane. The beam expands with the distance z from the waist according to

$$w^2(z) = w_0^2 \left[1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2 \right]$$

The resulting beam contour $w(z)$ is illustrated in the Figure 3. The far field diffraction angle θ is given by

$$\theta = \lambda / \pi w_0$$

The radius of curvature R of the wavefront changes according to

$$R(z) = z \left[1 + (\pi w_0^2 / \lambda z)^2 \right]$$

As the beam propagates, R varies from infinity at the waist to a minimum of

$$R = 2\pi w_0^2 / \lambda$$

at

$$z = \pi w_0^2 / \lambda$$

and approaches z asymptotically. Also, the beam experiences a phase shift θ relative to an ideal uniform plane wave given by

$$\theta = \tan^{-1} (\lambda z / \pi w_0^2)$$

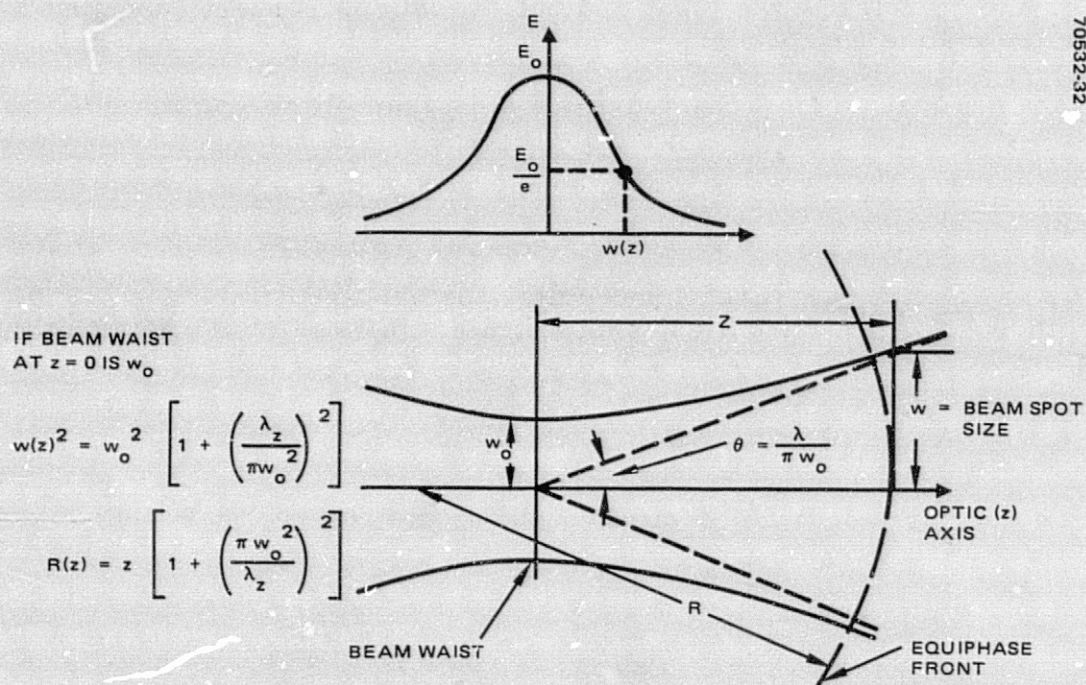


FIGURE 3. GAUSSIAN BEAM PARAMETERS

3) Complex Beam Parameter. The real beam parameters w and R are used to define a complex beam parameter q by

$$\frac{1}{q} = \frac{1}{R} - j \frac{\lambda}{\pi w_0^2}$$

Using the q parameter the laws of beam propagation simplify to

$$q_2 = q_1 + z$$

where q_1 and q_2 refer to the input and output planes, respectively, and z is the distance between these planes.

A thin lens of focal length f transforms the q parameters according to

$$\frac{1}{q_2} = \frac{1}{q_1} - \frac{1}{f}$$

where q_1 and q_2 are measured immediately before and after the lens. A beam passing through a general optical system characterized by its ABCD matrix are related by

$$q_2 = (Aq_1 + B)/(Cq_1 + D)$$

a solution which is accurate to within a phase factor

$$\theta \leq -j \ln \left(1 - \frac{\lambda z}{\pi w_0^2} \right).$$

Because of the combination properties of ray matrices described earlier, a complicated optical system of s elements is expressed simply as

$$M_s = M_1 M_2 M_3 \dots M_s$$

A beam waveguide consisting of a series of lens elements represented by the individual matrix M_L is simply

$$M_{\text{system}} = [M_L]^i$$

where there a total of i lenses in the system.

4) Beam "Smith" Charts. The complex beam parameter can be represented in a beam "Smith" chart in much the same way as a complex impedance parameter. Where the beam parameter is given by $1/q = 1/R - i2/b$, the complex plane of $i/q = i/R + 2/b$ is actually plotted. The imaginary axis is a plot of i/R while the real axis is a plot of $2/b$, where b is defined as the confocal parameter, $b = 2\pi w_0^2/\lambda$.

Any beam propagating in free space can be described by its radius of curvature at some point along the propagation axis and by its confocal parameter. Any propagating beam mode can be matched to any other through an appropriate set of interfaces, such as lenses or reflectors. The process of mode matching assures unconditional stability of the system, which means that the only diffraction losses suffered are those introduced by the limited aperture, or truncation loss of the Gaussian beam. Figure 4 illustrates the use of the complex beam parameter chart to design a quasi-optical feed for a radiometer. In Figure 4, there are three modes described by the appropriate confocal parameter for each mode, b_1 , b_2 , and b_3 .

The b_1 mode is determined by the characteristics of the antenna and is represented on the beam chart by a circle of radius $1/b_1$. It is instructive to "walk through" the system along the z -axis and to observe the techniques of mode matching.

- Enter Chart at $z = -\infty$. The point of beginning is from the antenna represented by $z_1 = -\infty$ on the chart. Moving along the z -axis in the positive direction is the same as moving along the circle of constant b past $z_1 = 0$ to the lens L_1 .
- Lens L_1 . The lens is the mode matching element to match the b_1 mode to the b_2 mode. We encounter the lens at approximately $z_1 = +1$ on the b_1 chart. The lens has the property of changing the radius of curvature of the wavefront. In this case, we want to change it from a positive value to a negative value suitable to match the b_2 mode. The length of line labeled L_1 on the chart tells us the power of the lens needed. Its focal length is $f = -1/L_1$.
- Jumping to Chart b_2 . In order to match modes, we must exit from chart b_1 at the same point on the imaginary axis as we enter chart b_2 . The beam chart representing region b_2 has a smaller radius than that of b_1 because the value of $1/b_2$ is smaller. We continue to move in the positive z direction until reaching lens L_2 .
- Lens L_2 . The second lens must have sufficient power to match the strongly converging beam mode b_3 , thus the power of the lens is greater than that of the first lens. Its focal length is given by $-1/L_2$.
- Jumping to Chart b_3 . Chart b_3 represents the converging beam which must match the feedhorn to the mixer. The confocal parameter b_3 is a small number because the beam waist is small and the radius of the circle in chart 3 is therefore large. We enter the chart from the matching lens at approximately the $z_3 = -1$ point and end at the feedhorn where the beam waist is smallest and the wave is a plane wave.

RECEIVING PAGE BLANK NOT FILMED

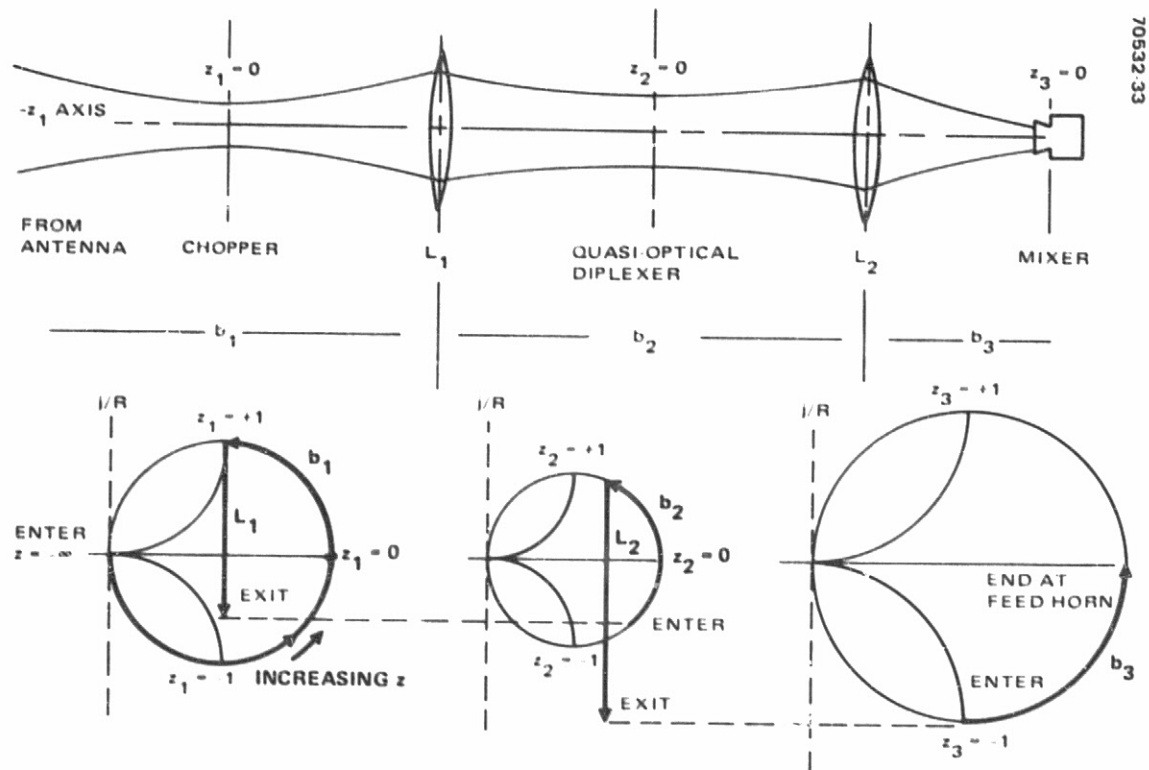


FIGURE 4. CIRCLE DIAGRAMS

3. Design Analysis

3.7 FABRY-PEROT FILTER DESIGN CONSIDERATIONS

The use of quasi-optical Fabry-Perot filters is anticipated for both local oscillator filters and for frequency multiplexing in MASR. Analytical techniques commonly used for optical systems are directly applicable.

The technique of generating multiple beam interference fringes from parallel plates was first reported by Fabry and Perot in 1899.⁽¹⁾ The technique has since been employed as a common spectroscopic tool in optics as the Fabry-Perot interferometer, consisting of plane parallel flat plates with reflective inner surfaces. When the two reflective surfaces are precisely parallel and are spaced at integral multiples of $\lambda/2$ of the wave passing through, an extremely sharp pass-band is produced. The sharpness of the bands make the tool uniquely useful as an optical spectral analyzer and as a quasi-optical local oscillator filter or multiplexer filter.

The following design relationships are extracted by the general analysis given by Born and Wolf.⁽²⁾ The relationships are given without proof for the purpose of aiding in the design application to MASR. The filter transmission characteristic is given by

$$P_{\text{out}}/P_{\text{in}} = \frac{1}{1 + F \sin^2\left(\frac{\pi}{2} \cdot \frac{f-f_0}{f_0}\right)}$$

and is shown in Figure 1. The reflection characteristic R is given by $R = P_{\text{ref}}/P_{\text{in}} = 1 - P_{\text{out}}/P_{\text{in}}$, where the F parameter is defined by

$$F = \frac{4R}{(1-R)^2}$$

The sharpness of the filter is enhanced for values of R approaching unity where the value of F becomes very large. The term $F \sin^2 [\pi/2 \cdot (f - f_0)/f_0]$ is zero for $f = f_0$, but becomes very large for values of $f \neq f_0$, thereby driving $P_{\text{out}}/P_{\text{in}}$ to zero.

Loss mechanisms in Fabry-Perot filters include ohmic losses, reactive losses due to poor reflective surfaces, and walk-off losses⁽³⁾ due to diffraction and geometric (tilting) (see Figure 1). Ohmic losses are estimated by the expression

$$L_{\text{ohmic}}(\text{dB}) = 8.6 \frac{Q_L}{Q_0}$$

where Q_L is the loaded Q of the resonator and Q_0 is the unloaded Q . The loaded Q is given by the ratio of signal frequency to bandwidth f_0/B_3 , while the unloaded Q is a function of the in-and-out coupling losses $Q_0 = \beta d/2a_c$. Unloaded Q s of Fabry-Perot resonators in the millimeter wave region range from 10,000 to more than 100,000.

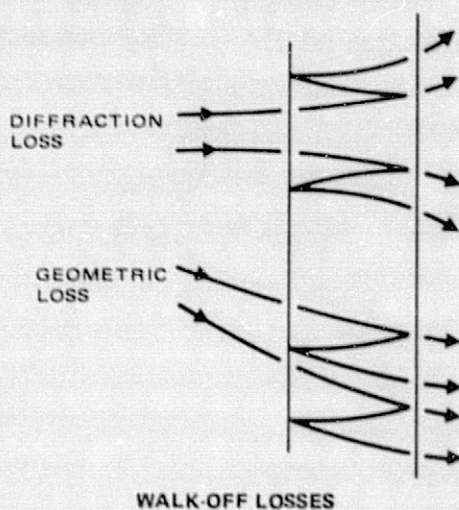
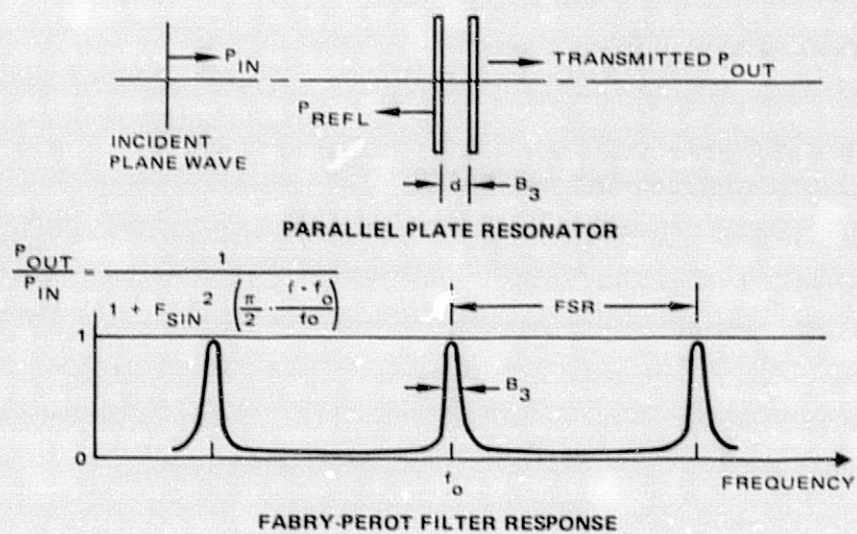


FIGURE 1. FABRY-PEROT PARALLEL PLATE FILTER CHARACTERISTICS

Finesse factor is defined by the ratio of free spectral range to the 3 dB bandwidth of the filter transmission band. The characteristic of the filter is such that the r th dB bandwidth is related to the 3 dB bandwidth by

$$B_r = B_3 (10^{r/10} - 1)^{1/2}$$

Table 1 lists these characteristics and other important design relationships of the Fabry-Perot filter for applications to MASR.

Practical application of Fabry-Perot filters to the millimeter wave region has been achieved by Goldsmith⁽⁴⁾ and Gustincic⁽⁵⁾.

References

1. C. Fabry and A. Perot, Ann. Chim. Phys. (7), 16, (1899), p. 115.
2. M. Born and E. Wolf, "Principles of Optics," Macmillan, New York, 1964, pp. 323-333.
3. J.A. Arnaud, A.M. Saleh, and J.T. Rushio, IEEE Transactions on MTT, Vol. MTT-22, No. 5, May 1974, pp. 486-493.
4. P.F. Goldsmith, "Quasi-Optical Feed for Millimeter Wave Radiometer," Technical Memorandum, Bell Telephone Laboratories, 11 January 1977.
5. J.J. Gustincic, "A Quasi-Optical Receiver Design," PGMTT Symposium, June 1977.

TABLE 1. DESIGN RELATIONSHIPS FOR QUASI-OPTICAL
FABRY-PEROT FILTERS

Finesse, $\mathcal{F} = \frac{\pi \sqrt{R}}{1-R} = \frac{\text{FSR}}{B_3}$	Reflectivity = R
Free spectral range, $\text{FSR} = \frac{c}{2d}$	3 dB bandwidth = B_3
10 dB bandwidth, $B_{10} = 3B_3 B_r = B_3 (10^{1/10} - 1)^{1/2}$	Plate spacing = d
Loaded Q, $Q_L = \frac{f_0}{B_3}$	Center frequency = f_0
Unloaded Q, $Q_0 = \frac{\beta d}{2\alpha_c} \left(\beta = \frac{2\pi}{\lambda} \right)$	Single pass coupling loss = α_c
Ohmic loss, $L = 8.6 \frac{Q_L}{Q_0}$	Midband rejection, $\text{dB} = 20 \log \left \frac{1+R}{1-R} \right $
Walk-off losses	
Diffraction parameter	$D = 2D/\beta\omega_0^2 (1-R)$
Geometric parameter	$G = 2d_{\sin \theta}/\omega_0 (1-R)$
For D and G ≥ 1 ,	Loss $\text{dB} \approx D + G$
For G $\ll 1$,	Loss $\text{dB} \approx 2.17R (1+R) G^2$

70532-351

4. RECEIVER DESIGN

4.1	MIXER DESIGN ALTERNATIVES	4-4
4.1.1	Waveguide Balanced Mixer	4-6
4.1.2	Quasi-Optical Balanced Mixer	4-8
4.1.3	Subharmonic Mixers	4-12
4.2	MIXER DIODES	4-16
4.2.1	EPI-Gallium Arsenide Material	4-16
4.2.2	Diode Design	4-18
4.3	LOCAL OSCILLATOR SOURCES	4-20
4.3.1	Transferred Electron Oscillators	4-20
4.3.2	IMPATT Sources	4-22
4.4	QUASI-OPTICAL LOCAL OSCILLATOR FILTERS	4-26
4.4.1	Comparison of Waveguide and Quasi-Optical Filters	4-26
4.4.2	Parallel Plate Versus Ring Resonator Configurations	4-28
4.4.3	Filter Design for 183 GHz Local Oscillator Filter	4-32
4.4.4	Designs for 118 GHz Local Oscillator Filter	4-33
4.4.5	Design for 104/140 GHz Local Oscillator Filters	4-35
4.5	IF DESIGN	4-36
4.5.1	Network for 0.625 to 10 GHz IF	4-36
4.5.2	Network for 20 to 4000 MHz IF	4-38
4.6	QUASI-OPTICAL DIPLEXERS	4-40
4.6.1	Separation of 183 GHz Band	4-40
4.6.2	Separation of 118 GHz Band From Window Bands Using Fabry-Perot Filter	4-44
4.6.3	Separation of 118 GHz Band From Window Band Using Resonant Grid Dichroic Filters	4-46

4. RECEIVER DESIGN

Where possible, recommended receiver component designs are described. In "soft" technology areas, such as mixers and local oscillator sources, various possible approaches are examined.

Mixer Design Alternatives. Three mixer designs look attractive for MASR. Selection of a preferred configuration from among these alternatives depends on development progress in waveguide balanced mixers, quasi-optical mixers, and subharmonically pumped mixers.

Mixer Diodes. High quality vapor phase epitaxial gallium arsenide material is difficult to obtain, and a batch selection and impound methodology is recommended for MASR flight programs. A conventional diode array of 1 to 2 μm dots, whisker contacted in a Sharpless wafer, is the recommended diode design.

Local Oscillator Sources. IMPATTs are primary contenders for use as LO sources in MASR. Quasi-optical filters used with balance mixers will provide sufficient LO noise rejection. The Gunn oscillator is unsuitable for use as a fundamental local oscillator due to a present high frequency limit of 100 GHz. Superior Gunn noise performance makes it an attractive alternative to the IMPATT when used with a doubler or with a subharmonic mixer.

Quasi-Optical Local Oscillator Filters. Fabrey-Perot quasi-optical filters have a much higher unloaded Q than waveguide filters, resulting in improved selectivity and lower loss.

IF Design. IF networks consisting of filters, duplexers, and amplifiers are described for the 118 and 183 GHz channels.

PRECEDING PAGE BLANK NOT FILMED

4. Receiver Design

4.1 MIXER DESIGN ALTERNATIVES

Three mixer designs look attractive for MASR. Selection of a preferred configuration from among these alternatives depends on development progress in waveguide balanced mixers, quasi-optical mixers, and subharmonically pumped mixers.

During the course of the present study, a number of mixer configurations were analyzed and evaluated; they include single-ended, waveguide hybrid balanced, crossbar balanced, quasi-optical balanced, and subharmonically pumped balanced configurations (Table I). The characteristics of each approach are summarized below.

The designs treated in detail in the succeeding topics of this subsection were selected as the most appropriate for the MASR application. Of those configurations that provide both reliability and high performance, the waveguide balanced mixer probably involves, at this time, the least technical risk. However, further development work in the subharmonic mixer may improve confidence and swing opinion in its favor.

Single-Ended. Single-ended mixers have been used more than any other type at these frequencies in ground based radiometers for radiotelescopes. The lowest reported conversion loss above 100 GHz has been achieved with a single-ended mixer. The disadvantages of this unit are difficulty in LO injection and extreme sensitivity to LO noise. It is estimated that 45 dB of LO filtering is required with IMPATT local oscillators. Filters having sufficient noise rejection at the IF frequencies also have high loss at the local oscillator frequency. Thus, the use of a single-ended mixer in the MASR system depends upon the availability of sufficient local oscillator power of the required spectral purity.

Waveguide Balanced Mixer. The classic waveguide balanced mixer using a magic T or a short-slot hybrid has been successfully used at 118 GHz and can be scaled up to 183 GHz. A balanced mixer has the advantages of local oscillator noise suppression, high local oscillator to RF isolation, and good IF source impedance. The main disadvantage is the added complexity due to the two mixer diodes per mixer. The use of a waveguide balanced mixer in MASR will be constrained to a top-wall hybrid or folded side-wall hybrid in order to provide contiguous LF outputs of the two diodes, minimizing the capacitance to help achieve a uniform impedance over the very wide IF bandwidths.

Crossbar Balanced Mixer. The crossbar balanced mixer has a good IF impedance characteristic that makes it attractive for use with very large IF bandwidths. The main disadvantages of this design are poor local oscillator to RF isolation, poor LO noise suppression, and poor reliability. The crossbar design has been used exclusively with beam-lead or pill-packaged diodes, limiting its use to the lower millimeter wave frequencies. Mechanical and thermal ruggedness of the crossbar mixer with whisker contact diodes remains to be proven.

Subharmonically Pumped Mixer. Subharmonic mixers have the distinct advantage of using a local oscillator at half the signal frequency. The excellent performance reported has been achieved at the lower millimeter wave frequencies using beam-lead diodes. Further development is required to prove that the subharmonic mixers can perform at the higher signal frequencies. The subharmonic mixers have the advantage of local oscillator noise cancellation.

Quasi-optical Balanced Mixer. The quasioptical balanced mixer is designed to use two single-ended waveguide mixers in a quasioptical hybrid assembly. This technique should be considered as a method of achieving local oscillator noise suppression without the usual concomitant losses.

References

1. G. T. Wrixon, "Low Noise Diodes and Mixers for the 1 to 2 mm Wavelength Region," IEEE Transactions MTT, Vol. MTT-22, No. 12, December 1974.
2. D. N. Held and A. R. Kerr, "Conversion Loss and Noise of Microwave and Millimeter-Wave Mixers," Parts I and II, to be published in IEEE Transactions MTT.
3. M. V. Schneider and G. T. Wrixon, "Development and Testing of a Receiver at 230 GHz," International Microwave Symposium Digest, June 1974.
4. A. R. Kerr, R. J. Mattauch, and J. A. Grange, "A New Mixer Design for 140 to 220 GHz," IEEE Transactions PGMTT, Vol. MTT-25, No. 5, May 1977.
5. A. R. Kerr, "Low Noise Room-Temperature and Cryogenic Mixers for 80 to 120 GHz," IEEE Transactions MTT, Vol. MTT-23, No. 10, October 1975.
6. J. J. Gustinic et al., "Extension of Schottky Diode Receivers into the Submillimeter Region," Symposium on Microwave Theory and Techniques, June 1977.
7. B. J. Clifton, "Schottky Barrier Diodes for Submillimeter Heterodyne Detection," IEEE Transactions MTT, Vol. MTT-25, No. 6, June 1977.
8. M. McColl, D. T. Hodges, and W. A. Garber, "Submillimeter-Wave Detection with Submicron-Size Schottky-Barrier Diodes," IEEE Transactions MTT, Vol. MTT-25, No. 6, June 1977.

TABLE 1. MIXER DESIGN ALTERNATIVES

Balanced Waveguide Mixers

- Topwall hybrid meets conditions of proven design and special low output capacitance needed for MASR IF bandwidth requirements.

Quasi-optical balanced mixer

- Quasioptical balanced mixer permits balanced effect to be achieved with two single-ended mixers.

Subharmonic mixers

- Subharmonic mixers offer good mixer performance with half local oscillator frequency. Local oscillator noise cancellation in IF is inherent in system.

70532-36T

4. Receiver Design

4.1 Mixer Design Alternatives

4.1.1 WAVEGUIDE BALANCED MIXER

Topwall or folded sidewall hybrids meet the requirement of local oscillator noise cancellation and the special requirement of low output capacitance for wide MASR IF bandwidths.

Balanced mixers have long been used to provide immunity to noise in the local oscillator and to provide convenient local oscillator injection.⁽¹⁾ Special wideband IF requirements of MASR are a special constraint for the output of the mixer. In the balanced mixer, output terminals should be joined and matched in the optimum characteristic impedance to the IF circuit. Physical distances between the mixer whiskers and the point where they are joined should be minimized. The waveguide hybrids that can meet this requirement are the topwall hybrid and the folded sidewall hybrid. In both of these structures, the IF outputs of the mixers are contiguous and can be conveniently matched to a transmission line to the IF circuit.

Figure 1 illustrates the topwall hybrid assembly as it may be used for the 183 GHz mixer. The signal enters through a scalar horn in one arm of the hybrid while the local oscillator enters another horn. After passing through the hybrid, the outputs are matched to an identical pair of mixer diodes mounted in Sharpless wafers^{2,3}. The wafers are shown separated from the interface flange, their respective outputs in close proximity. The IF outputs are individually filtered in the Sharpless mounts through an inverted LP coax filter. When mounted to the interface flange, the IF outputs of the individual wafers are joined through a coax feedthrough that has a characteristic impedance of about 100 ohms. The 100 ohm output is matched to the IF network, which may be a wideband amplifier (buffer) or a multiplexer filter network.

The advantages of the design are 1) a standard design that has been used at the lower microwave frequencies for many years; 2) identical diode waves — a feature that facilitates the testing, evaluating, and selecting of matched pairs; 3) optimum MIC geometry from mixers to IF; 4) demountable mixer diodes for easy replacement.

References

1. R. V. Pound, Microwave Mixers, Radiation Laboratory Series, Vol. 16, McGraw Hill, New York, 1948, pp. 257-287.
2. W. M. Sharpless, "Wafer-Type Millimeter Wave Rectifiers," BSTJ 35, November 1956, pp. 1385-1402.
3. W. M. Sharpless, "Point Contact Wafer Diodes for Use in the 90 to 140 GHz Frequency Range," BSTJ, September 1963, pp. 2497, 2498.

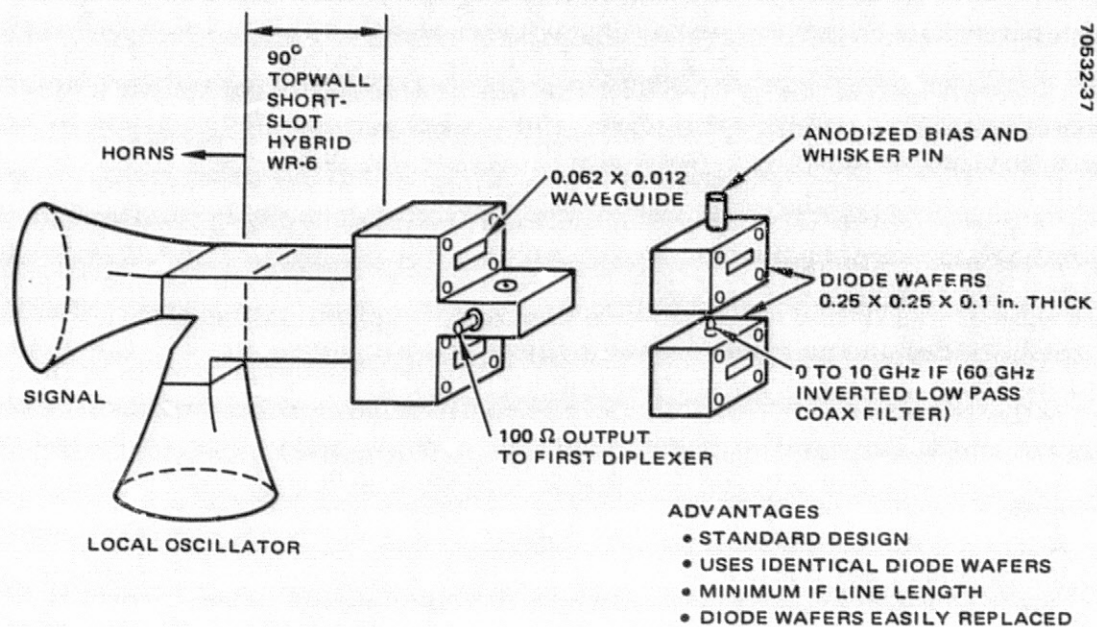


FIGURE 1. TOPWALL HYBRID ASSEMBLY FOR 183 GHz MIXER

4. Receiver Design

4.1 Mixer Design Alternatives

4.1.2 QUASI-OPTICAL BALANCED MIXER

The quasi-optical balanced mixer permits local oscillator noise cancellation to be achieved with two single-ended mixers and a quasi-optical hybrid assembly.

The quasi-optical balanced mixer using two single-ended mixers is made possible by combining the received signal and local oscillator through a polarization diplexer. Both signal and local oscillator (combined) signals are then divided with a second polarization splitter and focussed onto separate mixer mounts in such a way that IF variations caused by local oscillator noise are out of phase at the mixer outputs and are thereby cancelled with these outputs combined.*

For radiometric applications, the incoming signal polarization is random. Thus, selecting either one of the two incoming polarizations results in no difference in the signal quality. As shown in Figure 1, the polarization perpendicular to the paper is selected and reflected toward the top of the figure while the rejected polarization is transmitted by the wire grid polarizer 1. At the plane of rotation, the signal polarization is rotated 45° . This results in two components, one into the plane of the paper and one to the right in the plane of the paper. The local oscillator signal, which is incident on polarizer No. 1 from the bottom of the figure, is orthogonal in polarization to the signal and is transmitted through polarizer with no (or little) attenuation. After the rotation plane, the local oscillator polarization is shown separated into two components, one out of the plane of the paper and one to the right in the plane of the paper. Components of the signal and local oscillator in the plane of the paper are transmitted through the second polarizer (No. 2), reflected off the mirror at the top to the upper mixer mount. Components of the signal and local oscillator perpendicular to the paper are reflected off polarizer 2 to the right, reflected off the mirror at the right and focussed to the other mixer mount.

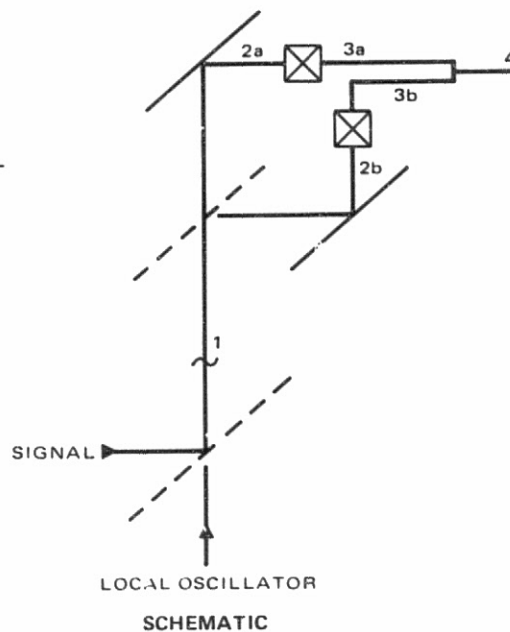
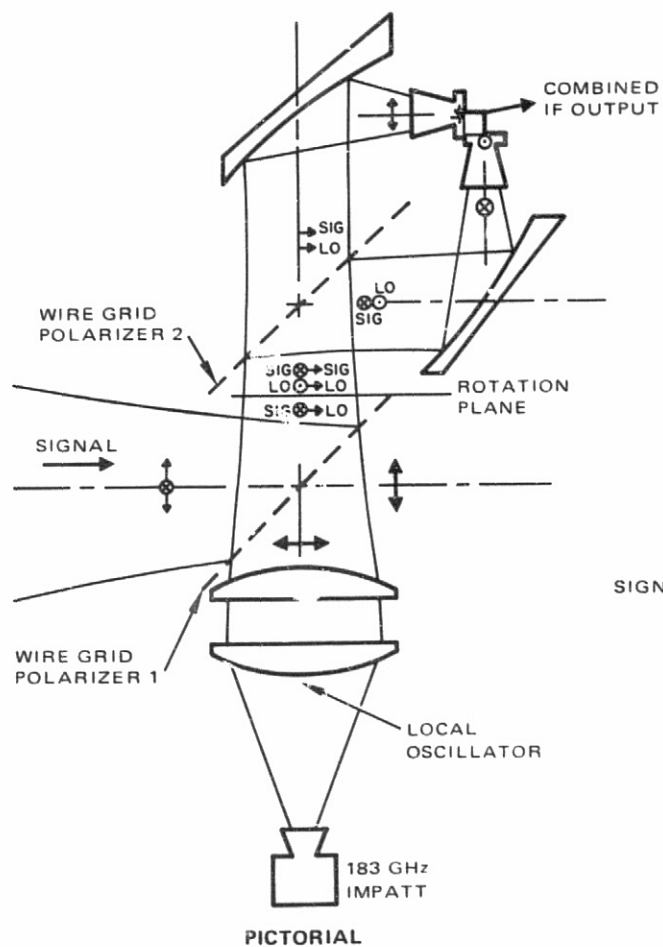
The signal and local oscillator are in phase in the upper mixer and are out of phase at the lower mixer. This results in the production of an IF signal at the upper whose noise components are out of phase with the IF signal at the lower mixer. Noise cancellation effects can be illustrated with a simplified analysis. Consider the simplified schematic diagram of the mixer in Figure 1. The signal and local oscillator fields can be represented at point 1 by

$$1) \quad E_s(t) = E_s \cos(\omega_s t + \phi_s) \quad P = -90^\circ \rightarrow -45^\circ$$

$$E_{LO}(t) = E_{LO} \cos(\omega_{LO} t + \phi_{LO}) \quad P = 0^\circ \rightarrow 45^\circ$$

The polarization of these fields with respect to the plane of the paper looking along the direction of propagation is indicated at the right. The changes caused by the polarization rotator are also indicated.

* This technique uses waveguide mounts for the mixer diodes and therefore differs in concept from the quasioptical single-ended mixer reported in J. J. Gustincic, "A Quasi-Optical Receiver Design," Symposium on Microwave Theory and Techniques, June 1977.



70532-38

FIGURE 1. QUASI-OPTICAL BALANCED MIXER

After passing through the second polarizer, both the signal and local oscillator fields are split evenly and directed to mixers a and b. The fields prior to entering the mixers at points 2a and 2b are

$$2a) \quad E_s(t) = \frac{E_s}{\sqrt{2}} \cos(\omega_s t + \phi_s + \beta_s a) \quad P = 0^\circ$$

$$E_{LO}(t) = \frac{E_{LO}}{\sqrt{2}} \cos(\omega_{LO} t + \phi_{LO} + \beta_{LO} a) \quad P = 0^\circ$$

$$2b) \quad E_s(t) = \frac{E_s}{\sqrt{2}} \cos(\omega_s t + \phi_s + \beta_s b) \quad P = -90^\circ$$

$$E_{LO}(t) = \frac{E_{LO}}{\sqrt{2}} \cos(\omega_{LO} t + \phi_{LO} + \beta_{LO} b + \pi) \quad P = +90^\circ$$

Upon entering the mixers, these fields produce voltage outputs which are proportional to the square of the vector sum of the signal and local oscillator fields

$$3a) \quad V_a = K \left[\frac{E_s^2}{2} + \frac{E_{LO}^2}{2} + 2 \frac{E_s}{\sqrt{2}} \frac{E_{LO}}{\sqrt{2}} \cos(\omega_{if} t + \phi_o) \right]$$

$$V_b = -K \left[\frac{E_s^2}{2} + \frac{E_{LO}^2}{2} + 2 \frac{E_s}{\sqrt{2}} \frac{E_{LO}}{\sqrt{2}} \cos(\omega_{if} t + \phi_o + \pi) \right]$$

which combined yield

$$4) \quad V_a + V_b = 2K E_s E_{LO} \cos(\omega_{if} t + \phi_o) \quad \phi_o = \phi_s - \phi_{LO}$$

The condition placed on this analysis is that $(\beta_s - \beta_{LO})(a - b) \ll \pi/2$, and since β_s and β_{LO} are determined by the local oscillator frequency and (highest) signal frequency, the maximum allowed difference in path length, a and b, can be estimated

$$(a - b) \ll \frac{c}{4(f_s - f_{LO})} \ll 7.5 \text{ mm}$$

This condition is easily achieved with state of the art mechanical alignment techniques.

4. Receiver Design

4.1 Mixer Design Alternatives

4.1.3 SUBHARMONIC MIXERS

Subharmonic mixers offer good mixer performance with half local oscillator frequency. Local oscillator noise cancellation in the IF is inherent in the system.

Subharmonic mixers using a local oscillator pump that is at half the normal LO frequency with an antiparallel diode pair can provide mixer conversion loss that closely approximates that of a normal mixer. In addition, this configuration provides local oscillator noise suppression. (1, 2)

The subharmonic mixer using half the normal LO frequency, $1/2 \times f_{LO}$, and an antiparallel diode pair can provide very efficient conversion of a signal from the frequency f_S to $|f_{LO} - f_S|$. This efficient mixing is accomplished because the LO power directly produces a time-varying conductance with a fundamental frequency variation at f_{LO} . This fact is demonstrated in Figure 1, which shows the conductance waveforms for a single diode and a diode pair pumped at $1/2 \times f_{LO}$. The diode pair has a time variation that is twice the pumping frequency. As in a standard mixer, the frequency conversion process in the subharmonic mixer occurs through the interaction of the signal power with the time-dependent conductance of the diodes. Since the primary frequency in the conductance waveform is f_{LO} , the principal mixing products are generated at $f_{LO} \pm f_S$, the same as in a standard mixer.

The signals generated at $f_{IF} = f_{LO} - f_S$ by the interaction of the LO noise with each diode have phases which cause them to cancel in the IF circuit. Imbalances in the diodes which create unequal conversion loss for the two diodes will degrade the noise cancellation. Likewise, these imbalances will cause the time-dependent conductance to have some half f_{LO} contribution. This will slightly degrade mixer performance, but should not be too serious in applications where sufficient LO power is available and where $f_{IF} \ll f_S$.

The question of how much LO power is needed has not yet been determined analytically. However, experiments by Bell Labs at millimeter waves have shown that diodes with low cutoff frequency or large parasitic elements require much higher pump power levels (~10 dB). Good mixer performance has been obtained with less than 10 dBm, while poorer diodes in the same circuit need 10 dB more LO power.

The best results have been obtained with diodes that have cutoff frequencies, f_{LO} , ~20x the signal frequency of the mixer. For the 180 GHz application, this implies that diodes with f_{LO} about 4000 GHz are needed. This figure of merit is probably beyond the present state of the art and can probably be met by utilizing the ultimate capabilities of GaAs materials technology and millimeter wave diode fabrication techniques.

Given the best diodes, it is still necessary to devise a mixer circuit with parasitics low enough not to degrade the device capabilities. The circuits that have performed best at millimeter waves have generally used some modification of the Sharpless wafer to contain the diode. The circuits that we propose for the subharmonic mixer also use that approach for embedding the diode in the RF circuit. Figure 2 shows two proposed mounting schemes for the diodes. Common to both approaches are the distributed low pass filters which complete the RF circuit for the diodes and allow IF power and bias currents to pass from the RF circuit region. With proper diode polarities, the close physical proximity of the diodes allows the IF

signals to be summed directly at the output of the modified wafer package. Minimizing the distance between the diodes and the IF amplifier has been one of the key elements in the past to achieving large IF bandwidths, and the Hughes design is consistent with this objective.

Figure 2 demonstrates two approaches to the RF circuitry. One (Figure 2a) uses an overmoded section of rectangular waveguide to pump the diodes at the subharmonic LO frequency, f_{LO} , and to couple the signal energy at f_S . The undesired mode that can propagate in the guide region is the TEU mode at f_S . If the circuit and device symmetries are maintained, this mode will not be excited. Any signal energy converted to this mode cannot propagate through the input transformer and will be rejected by the local oscillator filter so, that in principal, the circuit could be tuned to reconvert this energy to the IF. However, since the interaction region for the mode conversion is small, proper circuit fabrication and diode selection should minimize the problem. The alternate configuration (Figure 2b) solves the multimode problem by using ridge waveguide to provide broadband, single mode operation, but it also reduces the size of the region available for mounting the diodes and increases the RF losses in that part of the circuit. At 180 GHz, where the fundamental guide is only 0.050 inch wide, the size factor might be significant.

Reference

1. M. Cohn, J. E. Dregendorf, and B. A. Newman, "Harmonic Mixing with an Antiparallel Diode Pair," *IEEE Transaction on MTT*, Vol. MTT-23, No. 8, August 1975, pp. 667-673.
2. T. F. McMaster, M. V. Schneider, and W. W. Snell, Jr., "Millimeter-Wave Receivers with Subharmonic Pump," *IEEE Transactions on MTT*, Vol. MTT-24, No. 12, December 1976.

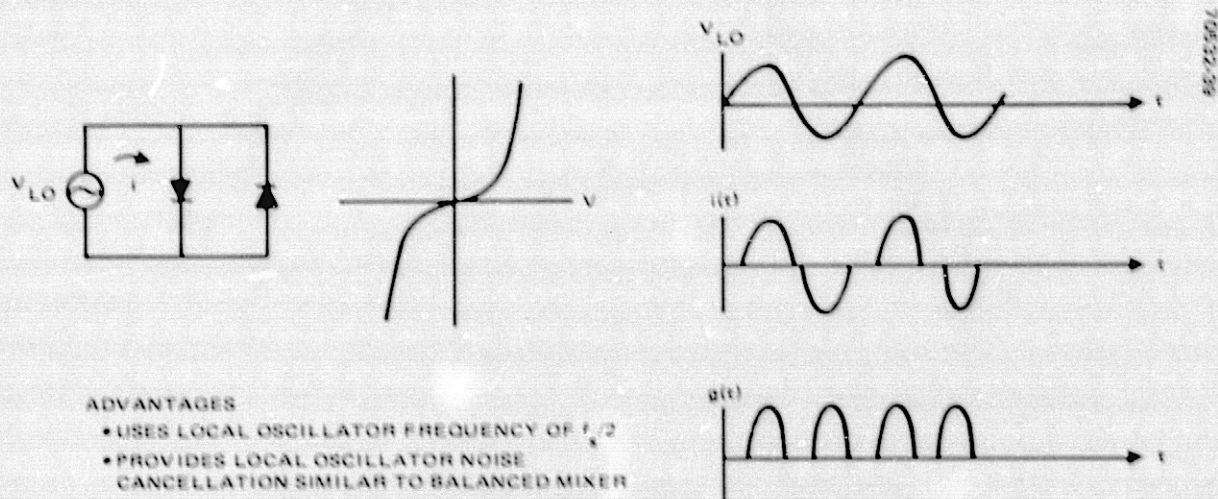


FIGURE 1. SUBHARMONICALLY PUMPED MIXER

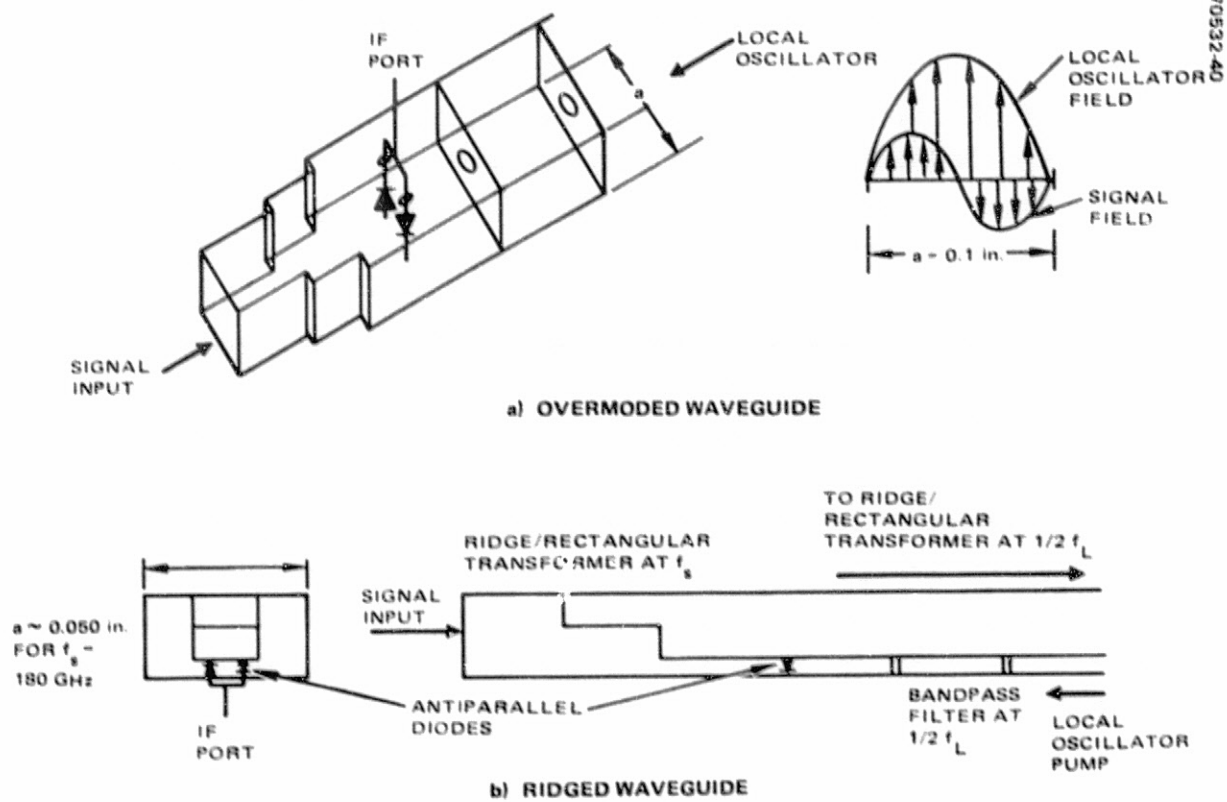


FIGURE 2. PROPOSED MOUNTING SCHEMES FOR DIODES

4.2.1 EPI-GALLIUM ARSENIDE MATERIAL

Mixer performance is critically dependent on high quality (vapor phase) epitaxial gallium arsenide material. Sources of this high quality material are extremely limited. Batch selection-and-impound methodology is recommended for MASR flight programs.

Performance of a mixer diode is dependent on geometry and upon the material. Candidate materials include silicon, gallium arsenide, and indium phosphide, in either a liquid phase or vapor phase epitaxial form. It is generally believed that the cutoff frequencies achievable with silicon are far below those achieved either with gallium arsenide or indium phosphide. Given the results reported to date, gallium arsenide appears to be the preferred material for frequencies from 100 to 200 GHz. Although indium phosphide shows promise of having even higher cutoff frequencies, its state of development is far behind that of gallium arsenide, and its selection as a candidate material involves greater risk.

The two types of gallium arsenide that have been considered are liquid phase and vapor phase epitaxial (VPE) material. Mixer conversion loss and noise figure are a critical function of the junction forward resistance, and whisker contacting of vapor phase material produces lower resistance and lower capacitance junctions. Some of the producers of vapor phase material are:

- Hughes Electron Dynamics Division (Torrance)
- Raytheon Research Center (Sudbury)
- GHz Devices (Cambridge)
- British Drug House (London)
- Plessey, Ltd. (London)

Production of quality VPE material suitable for low noise mixers is at present achieved accidentally, since full understanding of the growth control process has not yet been achieved. However, most of the above laboratories are currently improving facilities to accommodate a strong interest in Gunn diodes, FET amplifiers, and microwave mixers at the lower frequencies. It is reasonable to expect that these facilities will be able to produce batches for evaluation for mixers in the 200 GHz region. Certainly, one identified batch of high quality material is sufficient to provide all the mixers needed for a MASR flight program.

Evaluation of the material is not a simple process; it involves fabrication of diodes, contacting, mounting, and testing of samples from each batch produced. The methodology of EPI gallium arsenide evaluation and selection is illustrated in Figure 1. The recommended approach is to procure batches from a number of sources, fabricate sample diodes, and evaluate each of the batches from each source. When the quality needed for MASR has been achieved, the high quality batch is impounded for the exclusive fabrication of mixers for the program.

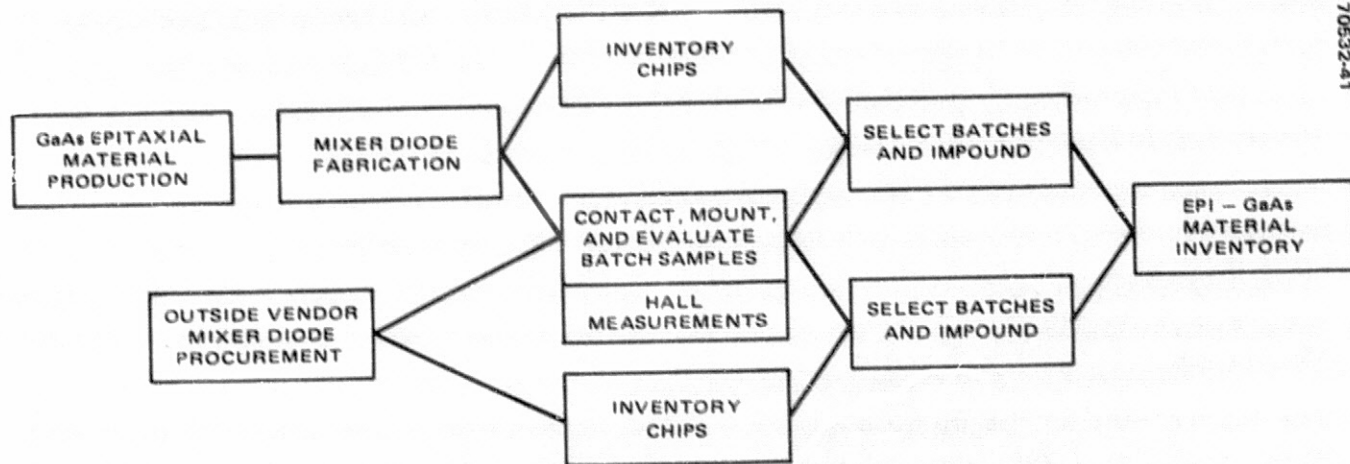


FIGURE 1. EPI GALLIUM ARSENIDE INVENTORY METHODOLOGY

4. Receiver Design

4.2 Mixer Diodes

4.2.2 DIODE DESIGN

A conventional honeycomb diode array of 1 to 2 μm dots, fabricated with electron beam lithography, is recommended. Sharpless wafers containing whisker contacted diodes are the most suitable packages for the 100 to 200 GHz region.

Photolithography and electron beam lithography are conventional methods of fabricating diodes from the basic EPI substrates. The smallest dots practical to achieve with photolithography are about 2 μm diameter, and some excellent diodes have been made by this method. Electron beam lithography is useful in producing sub- μm dots of great precision as well as exotic geometries such as the cross stripe geometry. These two configurations are shown in Figure 1. The cross stripe device has the advantages of lower series resistance and shunt capacitance and thus produces a higher cutoff frequency than the conventional dot geometry.⁽¹⁾ However, the simpler dot geometry of $\sim 2 \mu\text{m}$ diameter is probably adequate for MASR applications.

Contacting of whiskers to the dots is accomplished under either high power optical microscopes or scanning electron microscopes (SEMs). The resolution of optical microscopes is marginal for achieving precise contacting, and yields of diodes produced in this fashion are very low. SEMs give superior resolution and permit accurate contacting but have the disadvantage of requiring a vacuum environment and the use of micromanipulators to achieve proper contacting.

The importance of the diode package is directly related to the parasitic elements in the diode RF equivalent circuit. For low conversion loss, the cutoff frequency, f_{co} , has to be at least ten times the operating frequency (greater for subharmonic mixers), and the series resistance should be as low as possible. The cutoff frequency is given by

$$f_{co} = (2\pi R_s C)^{-1}$$

where

R_s = forward series resistance (8 ohms for best mixers) and

C = junction capacitance (0.007 pF for best mixers).⁽²⁾

The calculated cutoff frequency is thus 2.8×10^{12} Hz, which should produce a suitable mixer at frequencies up to 300 GHz.

A critical design element of the wafer mount is the IF output terminal. Figure 1 illustrates a typical mount with contacted diode chip and output IF pin. DC bias can be provided to the whisker through the insulated anodized whisker post, while local oscillator filtering in the IF is provided by the inside-out low pass filter built into the IF pin.

Reference

1. G. T. Wrixon, "Low Noise Diodes for Mixers for the 1 to 2 mm Wavelength Region, IEEE Transactions on MTT, Vol. MTT-22, No. 12, December 1974.
2. A. R. Kerr, R. J. Mattauch, and J. A. Grange, "A new Mixer for 140 to 220 GHz, IEEE Transactions, Vol. MTT-25, No. 5, May 1977.

70532-42

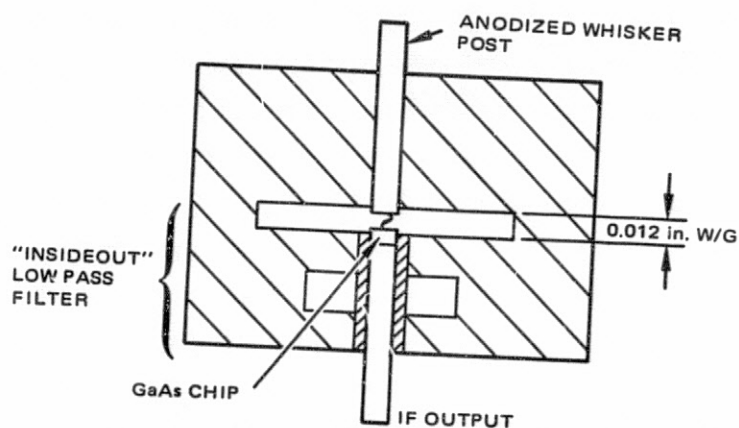
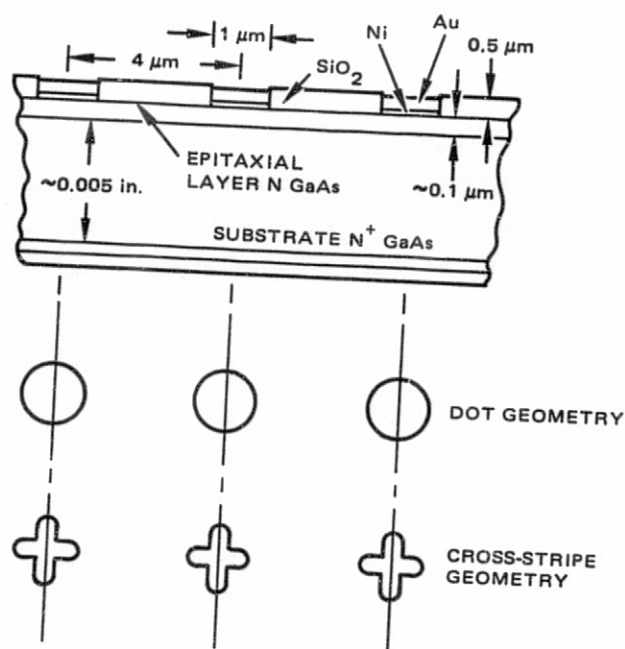


FIGURE 1. MIXER DIODE GEOMETRY/DIODE MOUNTING

4. Receiver Design

4.3 Local Oscillator Sources

4.3.1 TRANSFERRED ELECTRON OSCILLATORS

The Gunn, or transferred electron oscillator, is unsuitable for use as a fundamental local oscillator because of a present high frequency limit of about 100 GHz. Superior Gunn noise performance makes it an attractive alternative to the IMPATT when used with a doubler or with a subharmonic mixer.

Materials. Interest in indium phosphide as a Gunn material is growing because it has properties superior to those of the more widely used gallium arsenide. InP shows promise of achieving higher conversion efficiencies, reduced fabrication tolerances, and improved noise performance.

Power. The power presently available from CW Gunn oscillators is shown in Figure 1. The dashed line shows the present GaAs limit at 100 GHz. InP devices may extend this limit to ~150 GHz.⁽¹⁾

Reliability. The reliability of Gunn devices is quite good. Extensive life testing by Varian has established that the MTBF at the 90 percent confidence level is over 222,000 hours; at the 50 percent level, it is over 500,000 hours.⁽¹⁾ Three 55 GHz sources included in the Nimbus 5 satellite have operated flawlessly in space.

Noise. The carrier noise properties of Gunn devices are superior to those of IMPATTs, as shown in Figure 2. A filtered Gunn-doubler used with a single-ended mixer is a viable alternative to the filtered IMPATT-balanced mixer at 183 GHz.

Phase lock techniques used with Gunns are similar to those used with IMPATTs (see Topic 4.3.2). Circuit design and productization procedures are discussed in the literature. (See References 1 through 4.)

References

1. R. E. Goldwasser, J. F. Caldwell, "Millimeter Wave Gunn Devices," presented at WESCON, Los Angeles, September 1974.
2. R. J. Hamilton, R. D. Fairman, S. I. Long, M. Omori, and F. B. Fank, "InP Gunn-Effect Devices for Millimeter-Wave Amplifiers and Oscillators," IEEE Transactions MTT, Vol. MTT-24, No. 11, November 1976.
3. N. B. Kramer, "Millimeter-Wave Semiconductor Devices," IEEE Transactions MTT, Vol. MTT-24, No. 11, November 1976.
4. T. G. Ruttan, "High Frequency Gunn Oscillators," IEEE Transactions MTT, February 1974.

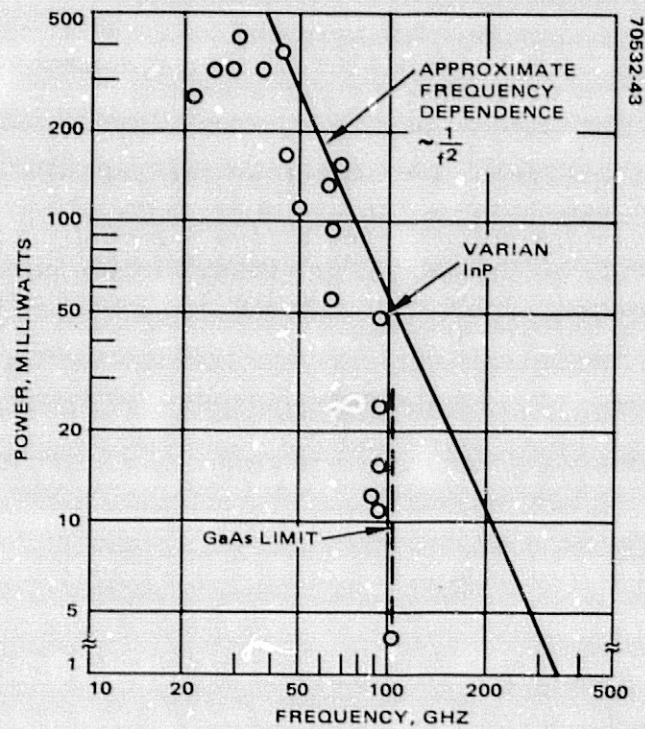


FIGURE 1. CW GUN OSCILLATOR POWER OUTPUT

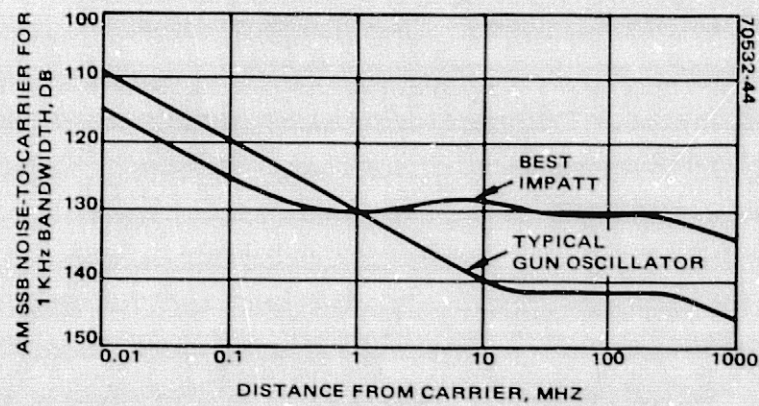


FIGURE 2. IMPATT/GUNN OSCILLATOR NOISE SPECTRA COMPARISONS

4. Receiver Design

4.3 Local Oscillator Sources

4.3.2 IMPATT SOURCES

Power, frequency range, and reliability of CW IMPATTs make them primary contenders for use as LO sources in MASR. Quasi-optical filters used with balanced mixers will provide sufficient LO noise rejection.

Power. Power output and frequency of state of the art IMPATT oscillators are shown in Figure 1. Adequate power is available in all channels using the 1977 state of the art and considering LO filter losses. A minimum available LO power of 6 mW occurs at 118 GHz because of high losses in the narrow LO filter that is required at that frequency. Increased 118 GHz LO power will become available as the IMPATT state of the art progresses.

Reliability. IMPATT lifetime is strongly dependent on junction operating temperature. Accelerated life testing of silicon millimeter-wave IMPATTs has indicated that the median time to failure is given by an Arrhenius relationship

$$\tau = \tau_0 \exp\left(\frac{E_a}{kT}\right)$$

where $\tau_0 = 1.6 \times 10^{-11}$ hours, the activation energy, $E_a = 1.6$ eV, and Boltzmann's constant, $k = 8.6 \times 10^{-5}$ eV/°K. This relationship predicts a median time to failure of 2×10^6 hours at 200°C.⁽¹⁾

Device defects are responsible for causing a higher short time failure rate than would be expected from the above relationship.⁽²⁾ Elimination of these defective devices by careful inspection during fabrication and an initial burnin will result in devices with reliability suitable for the MASR program.

Noise. A typical low Q IMPATT source generates AM noise which is significant out to high IF frequencies. The measured noise spectrum of a 60 GHz IMPATT (Figure 2) is reasonably flat to beyond 4 GHz, which implies a noise spectrum flat beyond 12 GHz for a 183 GHz IMPATT. The noise also varies with circuit parameters and bias current.

LO noise rejection of 40 dB over the IF bandwidth is desirable to eliminate the LO noise contribution to system temperature; 20 dB noise rejection is provided by the balanced mixer configuration, so that 20 dB rejection at the IF band edge is required from the quasioptical LO filters. This large rejection requires very high Q filters, especially at 118 GHz because of the 20 MHz IF band edge present. Quasi-optical Fabry-Perot filters were chosen over cavity filters because of their much lower losses.

Stability. The frequency drift temperature coefficient for a 118 GHz IMPATT (cavity Q_{ext} = 100) is expected to be in the range of 6 to 12 MHz/°C.⁽³⁾ In the absence of an accurate temperature controller, frequency stabilization will be required at 118 GHz to assure that frequency drift is less than one-fifth (lowest IF frequency). At 183 GHz, with a scaled temperature coefficient of 9 to 18 MHz/°C and a maximum allowable drift of 125 MHz, stabilization is desirable. The window channels at 104 and 140 GHz will not require stabilization.

Day-to-day IMPATT stability is good, but data on long term frequency drift due to aging is not available.

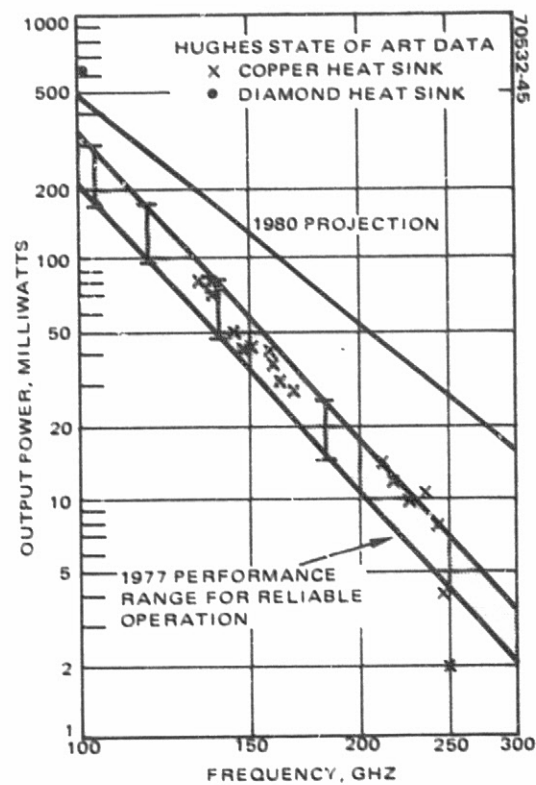


FIGURE 1. CW IMPATT OSCILLATOR POWER OUTPUT VERSUS FREQUENCY

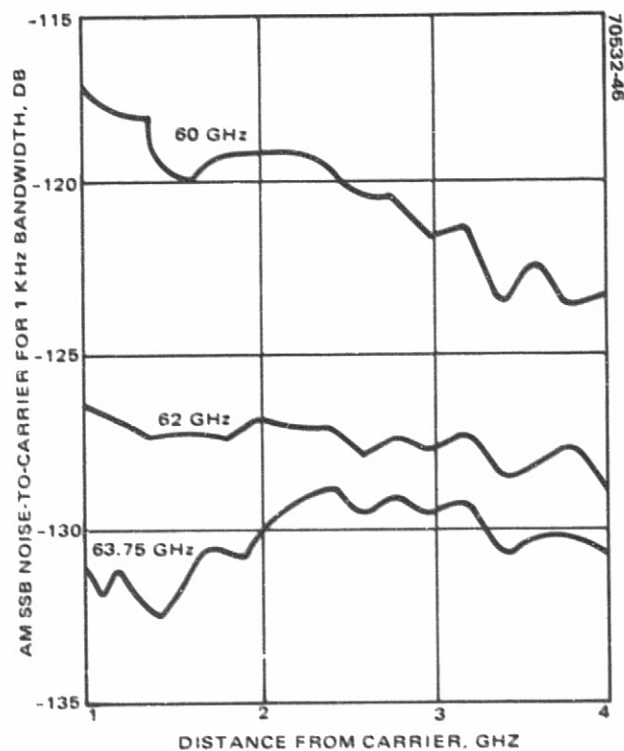


FIGURE 2. AM NOISE AWAY FROM CARRIER

Oscillator Circuit. Figure 3 is a cross section of a Y-band (170 to 260 GHz) IMPATT oscillator developed at Hughes.⁽⁴⁾ It consists of three major sections: a tapered waveguide section that transforms full height waveguide to reduced height, a reduced height waveguide wafer section that contains the IMPATT diode, and a mechanical tuning short section. The diode mount resembles the Sharpless wafer mount used in mixer diode applications. Use of the wafer mount facilitates the variation of circuit characteristics at the chip level, allowing the development of a diode package with optimum parasitics and good thermal properties. A detailed description of millimeter-wave IMPATT design and development procedures is given in the references listed below.

Phase Lock System. The IMPATT local oscillators will be phase locked to a stable reference to assure long-term frequency stability.⁽⁵⁾ A block diagram of the proposed phase lock system is shown in Figure 4. A sample of the IMPATT output is mixed with a harmonic of a lower frequency stabilized Gunn source, generating an IF signal close to the crystal reference. This IF signal is then locked to the reference by varying the IMPATT bias current, resulting in a phase locked IMPATT frequency of $f_S = Nf_{LO} \pm f_R$.

Locking to the proper sideband, above or below the nth harmonic of the Gunn, is assured by the use of two phase detectors working in quadrature. The output of the quadrature phase detector is sideband dependent, and is used with the search oscillator to move the IMPATT toward the desired sideband.

References

1. N. B. Kramer, "Millimeter-Wave Semiconductor Devices," IEEE Transactions MTT, Vol. MTT-24, No. 11, November 1976.
2. Peter Staeker, "K_a-Band IMPATT Diode Reliability," presented at the International Conference on Electron Devices, Washington, DC, December 1973.
3. K. P. Weller, private communication.
4. Chente Chao, Robert L. Bernick, Edward M. Makaji, Robert S. Ying, Kenneth P. Weller, and Don H. Lee, "Y-Band (170 to 260 GHz) Tunable CW IMPATT Diode Oscillators," submitted for publication in IEEE Transactions MTT 1977 Symposium issue.
5. J. Baprawski, C. Smith, and F. J. Bernues, "Phase-Locked Solid State Millimeter-Wave Sources," Microwave Journal, October 1976.
6. N. B. Kramer, "Millimeter IMPATT Devices," WESCON, Los Angeles, September 1974.
7. K. P. Weller, R. S. Ying, D. H. Lee, "Millimeter IMPATT Sources for the 130 to 170 GHz Range," IEEE Transactions MTT, Vol. MTT-24, No. 11, November 1976.

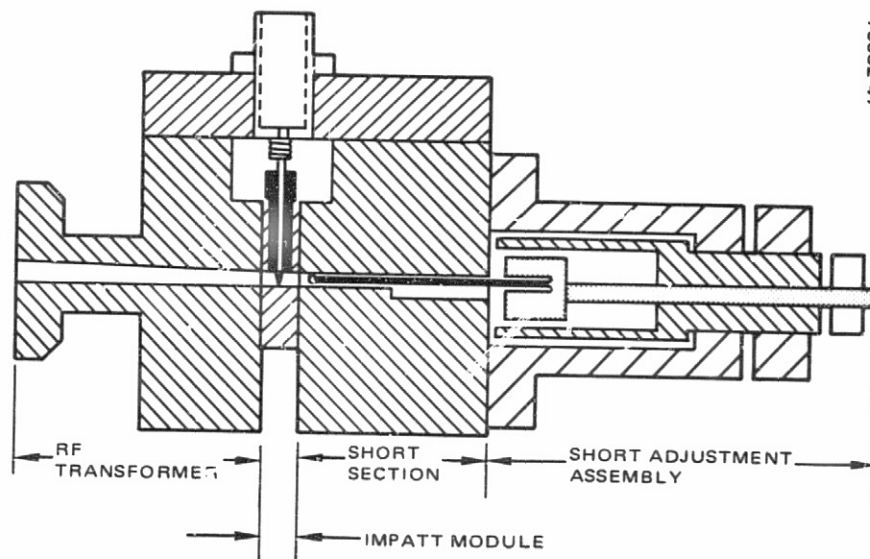


FIGURE 3. CROSS SECTION OF Y BAND IMPATT DIODE OSCILLATOR CIRCUIT

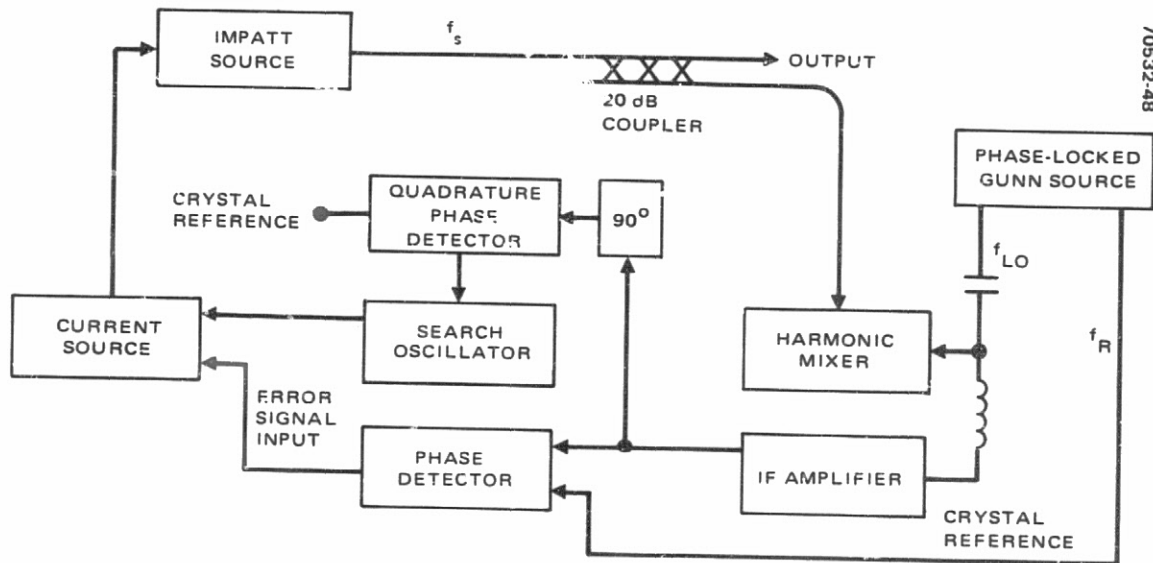


FIGURE 4. PHASE LOCKED MILLIMETER IMPATT SOURCE

4.4.1 COMPARISON OF WAVEGUIDE AND QUASI-OPTICAL FILTERS

Fabry-Perot quasi-optical filters have fundamental advantages over waveguide filters in relative high unloaded Q and selectivity for suppression of local oscillator noise in the IF sidebands.

The most critical requirement for a local oscillator filter is maximum suppression of out-of-band noise and minimum attenuation at the local oscillator frequency. From filter theory, the midband ohmic loss of a single pole filter is given by

$$\text{Loss (dB)} = 8.7 Q_L / Q_O$$

where Q_L is the loaded Q, f_o/B , and Q_O is the unloaded Q. Thus, in comparing waveguide filters with quasioptical filters, the first comparison we must make is of the limitations of unloaded Q, Q_O . The unloaded Q of a waveguide resonator is $Q_O = \pi Z_g / 4R_s$ where Z_g is the impedance of the waveguide and R_s is the surface resistivity of the metal. Typical Q_O values for TE₀₁₁ cavities range from 2000 at 100 GHz to 1500 at 200 GHz. However, the unloaded Q of a Fabry-Perot resonator is given by $Q_O = \beta d / 2\alpha_c$ where $2\alpha_c$ is the loss per pass of the wave into and out of the filter and βd is the enclosed phase factor of the filter in radians. Typical Q_O values for a Fabry-Perot filter at 200 GHz range from 10,000 to 25,000. It can be seen that the overall ohmic loss of the waveguide single pole device is ten or more times greater than that of the Fabry-Perot.

Required loaded Q's for local oscillator filters range as high as 6000. Since the circuit Q of a single pole waveguide filter is limited to less than 2000, a number of elements N must be used to improve the sharpness of the response. While the sharpness can be improved with a multipole filter, the insertion loss increases proportional to the number of sections.

A Fabry-Perot, on the other hand, can achieve circuit Q's of 6000 or more with a single pole, formed by the two mirrors of the resonator. The Fabry-Perot makes an ideal filter for local oscillators, since the out-of-band rejection increases with the circuit Q, and rejection values of 25 dB are easily achieved.

Table 1 lists ohmic insertion loss values for typical designs for waveguide and Fabry-Perot filters for the four radiometric bands. Also listed are actual performance data achieved by Goldsmith(1) and Gustincic(2) for parallel plate and ring resonator structures, respectively. It should be noted that the quasi-optical filter achieves approximately ten times less loss than does the waveguide device.

References

1. P. F. Goldsmith, "Quasioptical Feed for Millimeter Wave Radiometer," technical memorandum, Bell Telephone Laboratories, 11 January 1977.
2. J. J. Gustincic, "A Quasi-optical Receiver Design," PGMTT Symposium, June 1977.

TABLE 1. OHMIC INSERTION LOSS VALUES
FOR TYPICAL FILTERS

f_o	Q_L	Q_O	Rejection, dB	Loss, dB
Waveguide Filters, N = 5 Poles				
183	305	1750	25	7.6
140	280	1900	25	6.4
118	6000	2000	---	---
104	208	2000	25	4.5
Quasioptical Filters, N = 1 Pole				
120*	200	10000	26	0.2
190**	1600	5000	20	3.0
183	479	> 10000	> 25	0.42
140	280	> 10000	> 25	0.2
118	6000	> 10000	> 25	8.1
104	208	> 10000	> 25	0.2

*Parallel plate — Goldsmith

**Ring resonator — Gustincic

4. Receiver Design

4.4 Quasi-optical Local Oscillator Filters

4.4.2 PARALLEL PLATE VERSUS RING RESONATOR CONFIGURATIONS

Parallel plate Fabry-Perot filters have the advantage of arbitrarily large free spectral range while the ring configuration offers a convenient way of combining signal and local oscillator. The large IF bandwidth requirements for MASR make the use of parallel plate configurations mandatory.

An important requirement for a local oscillator filter is that the filter passbands have sufficient spectral separation to avoid noise bands in the IF of the receiver. For the individual MASR bands, the required spectral separations are:

183 GHz band, maximum IF frequency 10 GHz

140 GHz band, maximum IF frequency 2 GHz

118 GHz band, maximum IF frequency 4.1 GHz

104 GHz band, maximum IF frequency 2 GHz

While one of the LO filter passbands must lie at the center frequency of each line, the first passbands on either side must be separated from the center frequency by at least the values listed above. In a Fabry-Perot filter, the separation of passbands is called the free spectral range, FSR, and is given by $FSR = c/2d$, where c is the velocity of light and d is the separation between the plates.⁽¹⁾ This relation yields the following values of maximum plate separation for the individual bands:

183 GHz band $d < 15$ mm

140 GHz band $d < 75$ mm

118 GHz band $d < 37$ mm

104 GHz band $d < 75$ mm

For a parallel plate Fabry-Perot configuration, plate separations of values as low as 2 to 3 mm are easily achieved, while a ring resonator must maintain plate separation determined by the aperture of the device.

Figure 1 shows a parallel plate Fabry-Perot filter with its mirror surfaces slightly concave to correct for diffraction. The effective plate separation, d , is very nearly the physical separation at the mirror centers. The application for MASR is for the local oscillator source to be connected to the horn labeled P_{in} while the P_{out} horn is mounted to the mixer. Since out of band energy is reflected back to the source, a slight tilt (typically 3°) is required to avoid affecting the performance of the LO source. An isolator may alternately be used where tilt losses are too severe. This type of configuration is applicable to any of the MASR bands.

The ring resonator with local oscillator injection and provision for simultaneous LO filtering are also shown in Figure 2 (see subsection 3.8). It can be seen that the resonator spacing, d , is related to the beam waist, w_0 , and that the free spectral range is determined by the single pass distance, d , through the resonator. The specific example for a 140 GHz configuration is shown where, for the MASR tentative design, $w_0 = 27$ mm. The plate separation is thus greater than 216 mm,

and the free spectral range FSR is less than 1.38 GHz. Since the IF frequencies for this band may extend to 2 GHz, the ring resonator is not suitable for the MASR application.

The unique properties of the ring resonator for local oscillator injection should not be overlooked.⁽²⁾ As the signal is incident on the filter, energy at or near the 140 GHz frequency enters the filter and is lost while energy on either side of 140 GHz is reflected into the 140 GHz mixer. From the local oscillator side, energy at precisely 140 GHz is transmitted through the filter to the mixer while local oscillator noise on either side of the 140 GHz frequency is reflected at 90°. This convenient property of ring resonators also prevents reflected energy from returning to the IMPATT, which may otherwise cause instabilities.

References

1. J. J. Gustincic, "A Quasi-Optical Receiver Design," Proceedings of Microwave Theory and Techniques Symposium, 1977.
2. J. Payne, private communication, NRAO, Tucson, Arizona, September 1977.
3. Martin and Puplett, IR Physics, Volume 10, pp. 105-109, 1969.
4. N. R. Erickson, "A Directional Filter Diplexer Using Optical Techniques," IEEE Transactions on MTT, Volume 25, No. 10, October 1977.

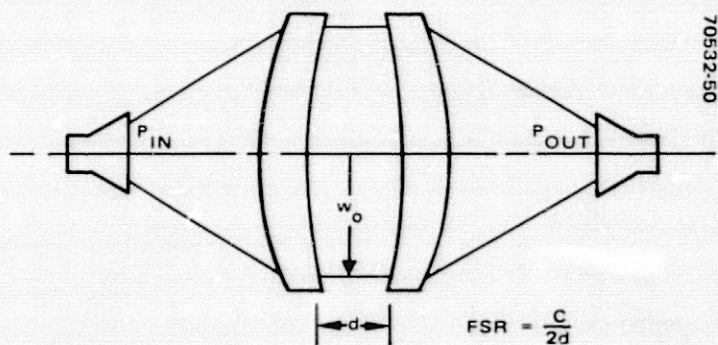


FIGURE 1. PARALLEL PLATE RESONATOR (GOLDSMITH)

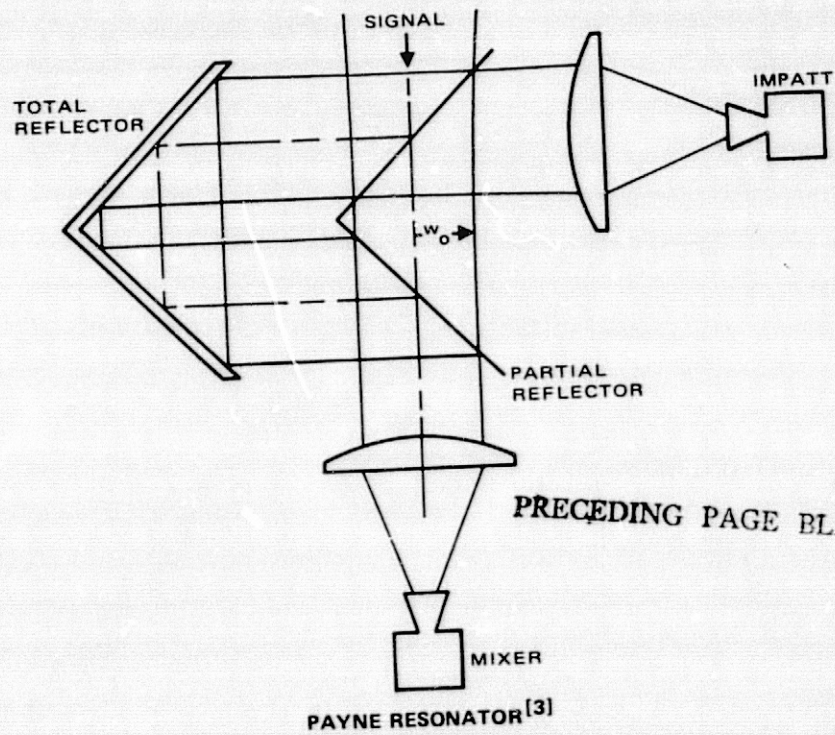
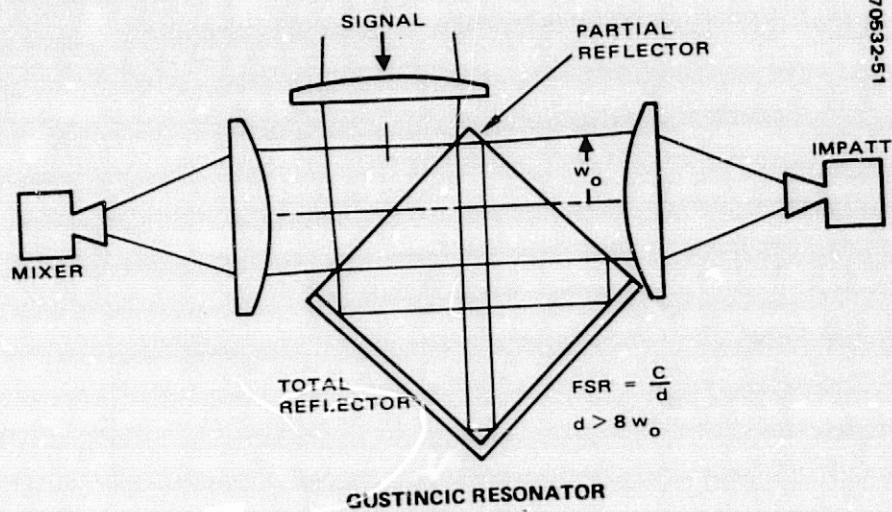


FIGURE 2. RING RESONATORS WITH LOCAL OSCILLATOR INJECTION

4. Receiver Design

4.4 Quasi-optical Local Oscillator Filters

4.4.3 FILTER DESIGN FOR 183 GHz LOCAL OSCILLATOR FILTER

The 183 GHz local oscillator filter has a total insertion loss of 1 dB and a midband rejection of more than 26 dB (Table 1) and Figure 1.

TABLE 1. 183 GHz LOCAL OSCILLATOR FILTER CHARACTERISTICS

Type	Fabry-Perot parallel plate
3 dB bandwidth	381 MHz
10 dB bandwidth	1200 MHz
Free spectral range	10 GHz
Beam waist	22.4 mm
Aperture diameter	70 mm
Walkoff diffraction loss	0.14 dB (without correction)
Reflectivity	0.91
Plate spacing	13 mm
Ohmic loss	0.42 dB
Midband rejection	26.5 dB
Walkoff geometric loss	0.4 dB for 3° tilt (no isolator used)

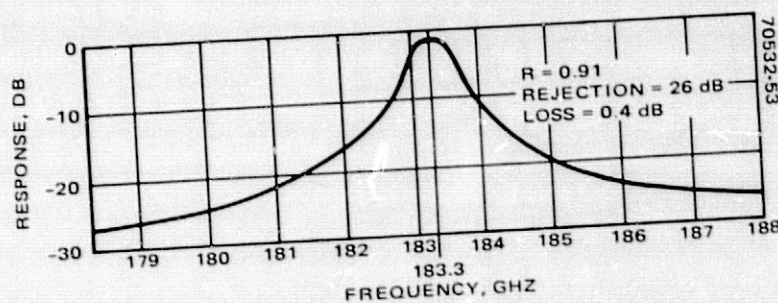


FIGURE 1. 183 GHz LOCAL OSCILLATOR FILTER RESPONSE

4. Receiver Design

4.4 Quasi-optical Local Oscillator Filters

4.4.4 DESIGNS FOR 118 GHz LOCAL OSCILLATOR FILTER

The most difficult filter for MASR application is for the 118 GHz local oscillator, where IF frequency of 20 MHz must be filtered. The largest contribution to loss is thus from ohmic loss (8.1 dB), since diffraction loss can be corrected for and tilt loss can be avoided by using an isolator. The filter characteristics and response are summarized in Table 1 and Figure 1.

TABLE 1. 118 GHz LOCAL OSCILLATOR FILTER CHARACTERISTICS

Type	Fabry-Perot parallel plate (and corrected plate)			
3 dB bandwidth	26 MHz			
10 dB bandwidth	40 MHz			
Free spectral range	4 GHz			
Beam waist	48.3 mm	without correction optics	22.4 mm	with correction optics
Aperture diameter	150 mm		70 mm	
Walkoff diffraction loss	1.3 dB		0 dB	
Reflectivity	0.99			
Plate spacing	38 mm			
Ohmic loss	8.1 dB			
Midband rejection	54 dB			
Walkoff geometric loss	8.2 dB for 3° tilt (isolator recommended rather than tilt)			
	4.9 dB for 1.8° tilt			
	0 dB for no tilt			

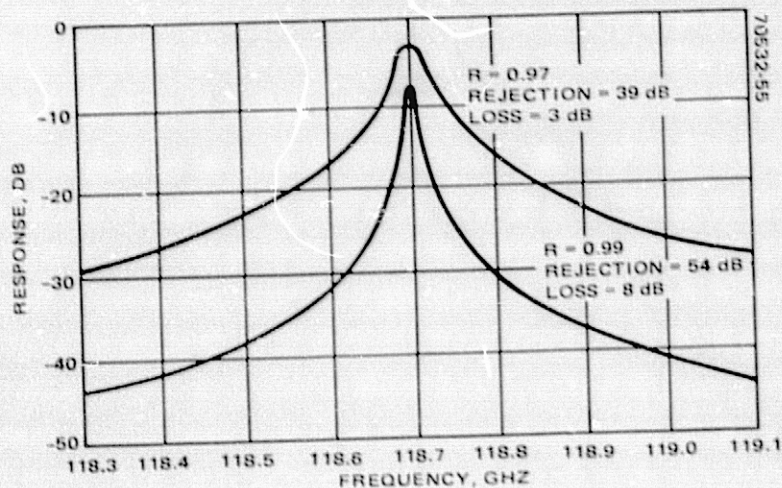


FIGURE 1. 118 GHz LOCAL OSCILLATOR FILTER RESPONSE

4. Receiver Design

4.4 Quasi-optical Local Oscillator Filters

4.4.5 DESIGN FOR 104/140 GHz LOCAL OSCILLATOR FILTERS

Filters for the 104 and 140 GHz window bands are of similar design. A midband insertion loss of 1.4 and 1.9 dB with walkoff losses of 4.6 dB is achievable with a midband rejection of 26 dB. Filter characteristics are summarized in Table 1.

TABLE 1. 104/140 GHz LOCAL OSCILLATOR FILTER

Type	Fabry-Perot parallel plate
3 dB bandwidth	64 MHz
10 dB bandwidth	200 MHz
Free spectral range	2 GHz
Beam waist	22.4 mm
Aperture diameter	70 mm
Walkoff diffraction loss	1.1 dB (without correction)
Reflectivity	0.89
Plate spacing	75 mm
Ohmic loss	1.4 dB at 104 GHz, 1.9 dB at 140 GHz
Midband rejection	25.6 dB
Walkoff geometric losses	3.5 dB for 3° tilt

70532-56T

~~PRECEDING~~ PAGE BLANK NOT FILMED

4. Receiver Design

4.5 IF Design

4.5.1 NETWORK FOR 0.625 TO 10 GHz IF

A network of contiguous diplexers, filters, and amplifiers can provide satisfactory IF multiplexing and noise figures for each of the six H_2O channels.

The state of the art of contiguous filter banks is progressing rapidly and has reached the point where it can be stated with assurance that a six-way evanescent mode contiguous multiplexer to meet the 0.625 to 10 GHz requirement is feasible*. Such a filter would have low loss and would present a substantially constant input impedance to the mixer over the entire band. However, the design procedures for such filters are new, and a custom design requires further effort beyond the scope of this report.

Conventional designs using contiguous diplexers in cascade are more common and are directly applicable to the problem. Such a design has been worked out as an example and is shown in Figures 1 and 2.

The critical part of the filter is the impedance match between the mixer and the first diplexer. From the output of the hybrid balanced mixer, the filter input coax connects to the junction of a high pass low pass diplexer. The nominal impedance of the coax is 100 ohms, the approximate output impedance of a balanced mixer. The impedance of each filter is also 100 ohms, increasing at the crossover frequency (5 GHz). At the crossover frequency, the shunt impedance of the two parallel branches is also 100 ohms.

The 5 to 10 GHz portion is again split into two bands in another evanescent mode contiguous diplexer whose outputs cross over at 7.5 GHz; each band feeds an IF amplifier. Similarly, the band from 0.6 to 5.0 GHz is also split into two bands which cross over at 2.5 GHz. Finally, the lower bands are each split again so that six bands are produced, each with a 50 ohm output and each with an IF amplifier. Figure 2 shows the complete 183 GHz band IF circuit with six output channels, identifying the band of frequencies covered by each channel and the respective IF noise figures.

Using the relationships developed in Section 3, and the effective IF noise figures, F_{if} are calculated which include the effects of filter losses. The radiometer ΔT 's are calculated for a conventional Dicke radiometer. These calculations assume a conservative 8.5 dB mixer conversion loss and an 0.5 dB feed loss. A typical antenna temperature of 250°K is also assumed. It should be noted that the calculated values of ΔT exceed the goals by about 1.7 times. These values may be improved by achieving a lower mixer conversion loss or by the use of reference averaging in the processing of data.

*For example, the Filtronics (Santa Clara, California) quintuplexer, 0.6 to 12 GHz 1.5 dB maximum loss, 55 dB isolation, 2.5 to 1 maximum VSWR.

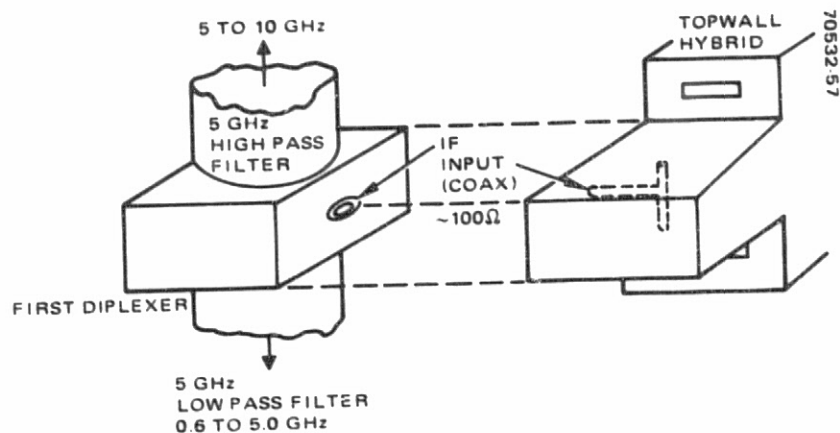


FIGURE 1. WAVEGUIDE BALANCED MIXER OUTPUT INTERFACE TO IF NETWORK

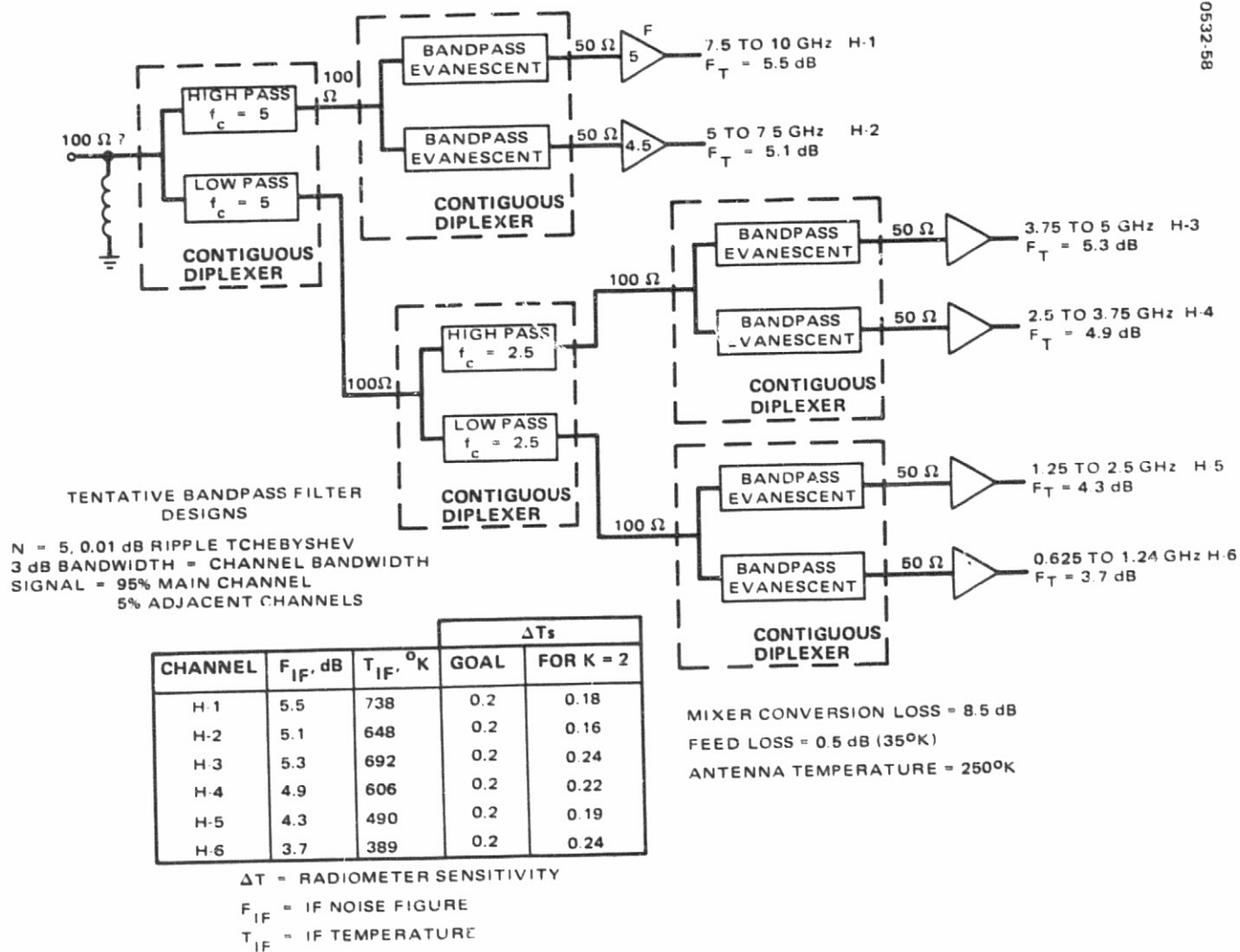


FIGURE 2. 183 GHz IF NETWORK DESIGN

4. Receiver Design

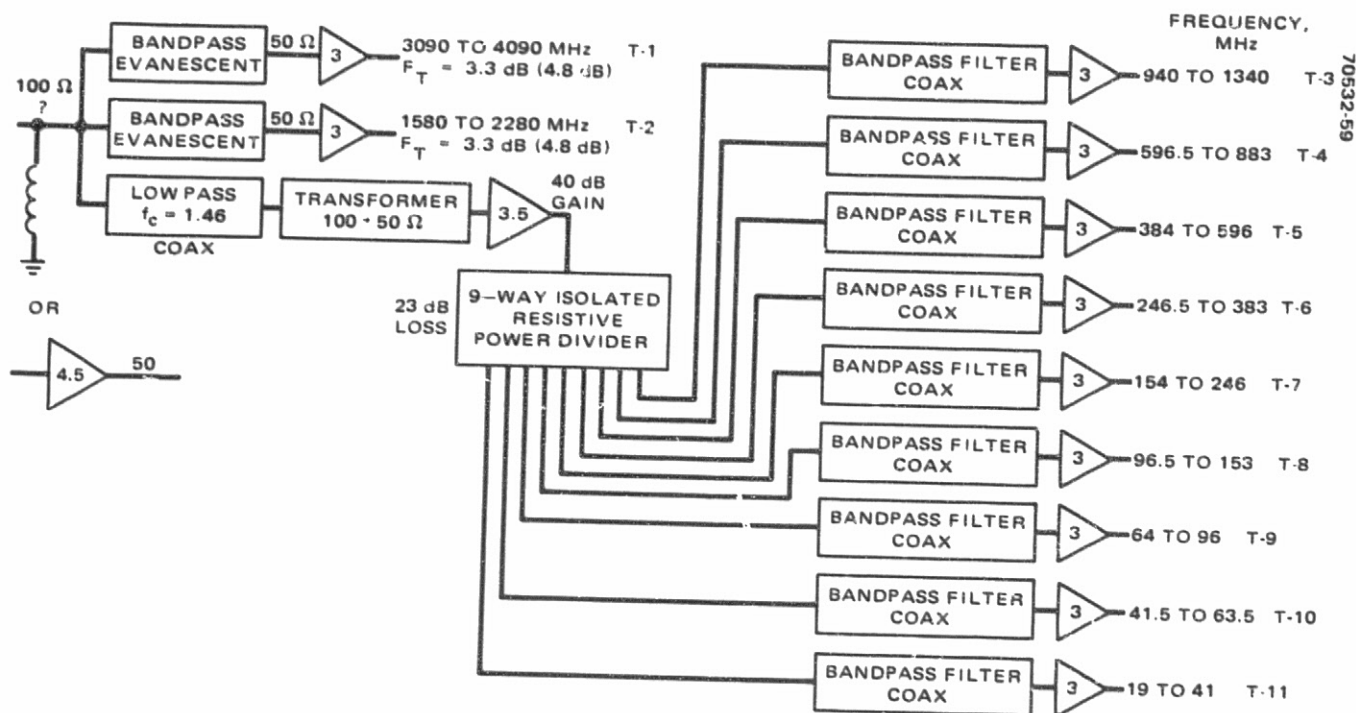
4.5 IF Design

4.5.2 NETWORK FOR 20 to 4000 MHz IF

Satisfactory IF multiplexing and noise figure performance for the 11 temperature channels can be provided by direct input to a filter network, but a broadband pre-amplifier between the mixer and filter could be beneficial to isolate the mixer from the filter network.

The optimum performance of a mixer is achieved with an optimum termination for the mixer at the IF frequency. Broadband GaAs FET devices are presently available with meet the need of a first stage IF amplifier from 20 to 4000 MHz. The following discussion treats two cases: 1) the use of a triplexer filter directly connected to the mixer and 2) use of a GaAs FET amplifier as a buffer between the mixer and filter. In Figure 1 the input to the triplexer is shown with a tentative 100 ohm input impedance. Directly below, an optional amplifier is shown. The noise figure of the amplifier at the high end of the band is 4.5 dB; at the low end it is about 2.5 dB. The signal enters the triplexer filter assembly and is divided into three bands: 3090 to 4090 MHz, 1580 to 2280 MHz, and 20 to 1340 MHz. The lower band is amplified with a 40 dB amplifier and enters a nine-way isolated resistive power divider having a 23 dB loss. The nine outputs are separately filtered with bandpass coax filters, amplified, and then detected. The effective noise figures of each channel F_{if} is given at the output of each amplifier. In parenthesis, a larger noise figure is indicated for the two high frequency channels. The noise figure in parenthesis is that seen when the preamplifier is inserted between the mixer and the triplexer; without the amplifier, the noise figure is determined by the 3.0 dB IF amplifiers for each of these channels plus the effects of losses introduced by the filter. However, when the amplifier is inserted, the noise figure is dominated by the front-end amplifier. Thus, the use of a broadband front-end amplifier is justified only as a buffer to aid in the broadband impedance match between the mixer and IF, and is achieved only at the expense of the noise figure of the higher frequency channels.

Calculations of radiometer ΔT performance for each of the T channels is given in Figure 1. The IF noise figures assumed are those obtained without the broadband amplifier. The mixer conversion loss assumed is 5 dB, a value which is considered within the state of the art at 118 GHz. An antenna feed loss of 1 dB and an antenna temperature of 250°K are assumed. It should be noted that the calculated ΔT values are lower than the values listed as goals. The conclusion of this assessment is that the 118 GHz receiver technology is well within the state of the art.



CHANNEL	F_{IF} , dB	T_{IF} , °K	ΔT_s	
			GOAL	FOR $K = 2$
T-1	3.3	330	0.2	0.65
T-2	3.3	330	0.2	0.08
T-3	3.5	359	0.2	0.11
T-4	3.5	359	0.2	0.14
T-5	3.5	359	0.2	0.16
T-6	3.0	290	0.2	0.18
T-7	3.0	290	0.2	0.22
T-8	3.0	290	0.2	0.27
T-9	2.5	225	0.2	0.34
T-10	2.5	225	0.2	0.40
T-11	2.5	225	0.2	0.40

MIXER CONVERSION LOSS = 5 dB
 FEED LOSS = 1 dB
 ANTENNA TEMPERATURE = 250°K

ΔT = RADIOMETER SENSITIVITY

F_{IF} = IF NOISE FIGURE

T_{IF} = IF TEMPERATURE

FIGURE 1. 118 GHz IF NETWORK DESIGN

4. Receiver Design

4.6 Quasi-optical Diplexers

4.6.1 SEPARATION OF 183 GHz BAND

Wire grid polarizers are the lowest loss elements capable of separating millimeter wave bands. It is logical to use such a device to separate the 183 GHz band from the other bands.

The most difficult radiometric band is the 183 GHz band. High mixer conversion loss, difficult IF, and local oscillator source noise and power limitations combine to produce a receiver noise temperature significantly higher than any of the other radiometric bands. Input feed losses add further to the input noise temperature. Since the 183 GHz band is a key MASR sounding band, one approach is to minimize input feed losses in this band even at the expense of increased losses in the other bands.

If it was desirable to minimize losses in the 118 GHz band at the expense of the 183 GHz band, the 118 GHz band would be separated using a polarizer and the 183 GHz band would be removed with a high pass filter, resulting in 0.5 dB feed loss at 118 GHz and 1.0 dB loss at 183 GHz. The resulting ΔT 's would be those presented in Topic 2.1, improved at 118 GHz and degraded at 183 GHz by a factor of 1.12

A wire grid polarizer functions as a diplexer in a radiometer because both polarizations are present in the signal input. Since a heterodyne mixer is sensitive only to one polarization, the other polarization is usually lost in a conventional receiver. Thus, a wire grid polarizer can be used to reflect the 183 GHz band of one polarization, and to transmit all other bands of the opposite polarization. Figure 1 shows the polarizer disposed at 45° with respect to the incoming signal, which contains both polarizations. The reflected polarization contains the entire band of frequencies, but only the 183 GHz portion is used. The transmitted polarization contains the entire band of frequencies, but only the lower bands are used.

Wire grid polarizers have long been in common use as microwave devices. Implementation is routine at the lower microwave frequencies but is not difficult even at 183 GHz. Two fabrication techniques are commonly used. The photo-etch method consists of photo-etched copper strips on a mylar or Kapton substrate. (1, 2) Figure 2 shows the design for a midband frequency of 150 GHz, with reactive and ohmic losses estimated for 183 GHz. Losses in the mylar or Kapton are not insignificant, but are kept low because the thickness of the sheet is only 50 μm (0.002 inch). Another fabrication method is by hand winding phosphor bronze wire on screws of very small pitch, 100 to 250 turns per inch. The tightly stretched grid of wires formed by the screw pitch is bonded to an elliptical ring, and the excess wire is cut away after the bond is set. The resulting polarizer will provide a circular aperture at an incident angle of 45° . Figure 3 illustrates winding the fine wire on the screws, bonding to the ring, and finishing the final polarizer. Performance is determined by the wire size, spacing, and wavelength, as shown in Figure 4.

References

1. T. S. Chu, M. J. Gans, W. E. Legg, "Quasi-optical Polarization Diplexing of Microwaves," BSTJ, Vol. 54, No. 10, December 1975.
2. J. Payne, private communication, National Radio Astronomy Observatory, Kitt Peak, Tucson, Arizona, September 1977.

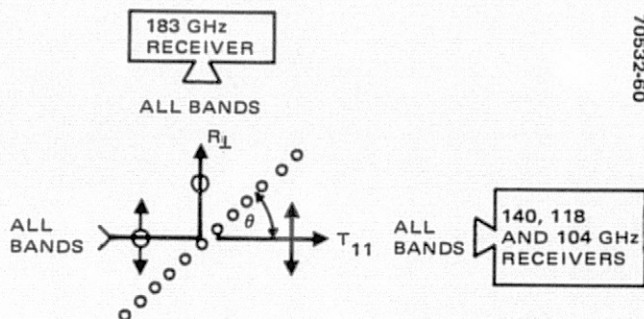
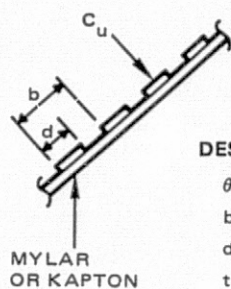


FIGURE 1. BASIC WIRE GRID POLARIZER

70532-60



DESIGN PARAMETERS

- $\theta = 45^\circ$ ANGLE OF INCIDENCE
- $b = 0.1$ mm GRATING PERIOD
- $d = 0.04$ mm COPPER STRIP WIDTH
- $t = 0.005$ mm COPPER STRIP THICKNESS
- $t' = 0.05$ mm MYLAR THICKNESS

REACTIVE LOSSES

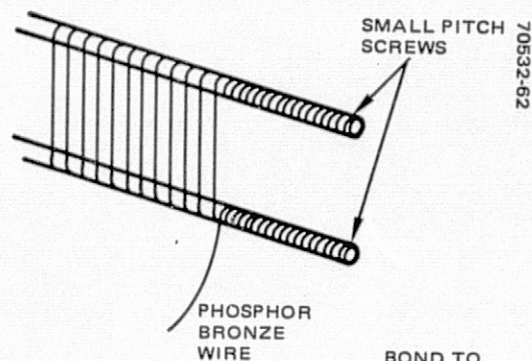
- $T_{11} \leq 0.001$ (0.004 dB)
- $R_1 \leq 0.001$ (0.004 dB)

OHMIC LOSSES

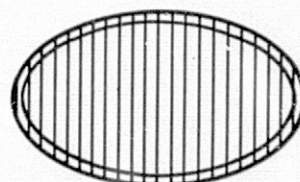
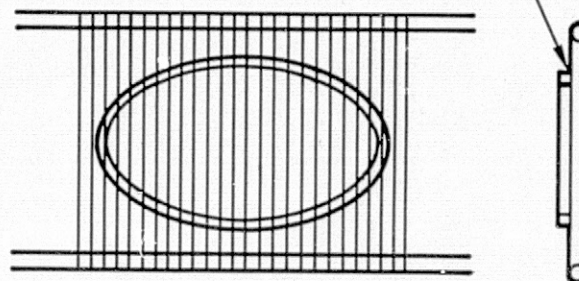
- $T \leq 0.003$ (0.01 dB)
- $R_1 \leq 0.001$ (0.004 dB)

FIGURE 2. PHOTOETCH FABRICATION METHOD

70532-61



70532-62



FINAL POLARIZER AFTER CUTTING AWAY WIRE

FIGURE 3. WIRE MESH FABRICATION METHOD

C-2

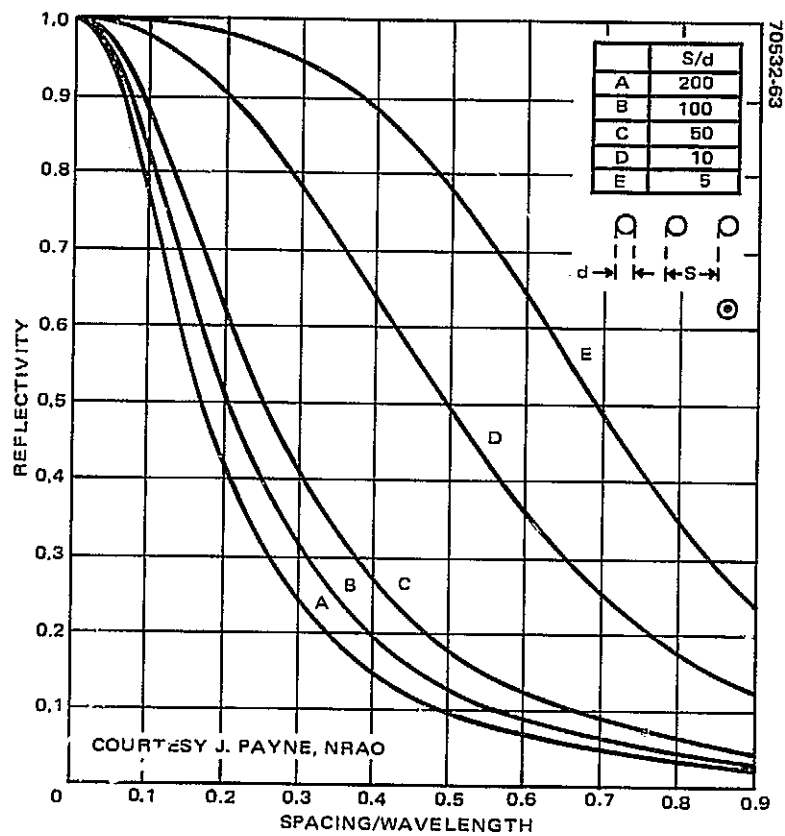


FIGURE 4. WIRE GRID POLARIZER DESIGN NOMOGRAPH

PRECEDING PAGE BLANK NOT FILMED

4. Receiver Design

4.6 Quasi-optical Diplexers

4.6.2 SEPARATION OF 118 GHz BAND FROM WINDOW BANDS USING FABRY-PEROT FILTER

A Fabry-Perot etalon offers a unique and low loss method of separating window bands from the 118 GHz band. However, the tilt angle must be kept small to minimize geometric walkoff losses.

The characteristics of Fabry-Perot filters are such that low loss, narrow transmission bands and broad reflection bands can be achieved. (1) For some applications, the Fabry-Perot is an ideal diplexer, but for other, such as for a high pass or low pass application, the device is poorly suited. (2 to 4) A unique application is seen for the Fabry-Perot in separating the broad 118 GHz band from the window bands at 104 and 140 GHz. The required bandwidths at 104 and 140 are about 4 GHz, permitting a loaded Q of about 30. Using the relations given in Subsection 3.8, a filter can be designed which has a transmission loss of only 0.1 dB and reflection loss of 0.14 dB. The layout of the filter is shown in Figure 1. The filter is tilted with the incident axis and the 118 GHz band is reflected at an angle of 30°. The unique property of the filter is that the third and fourth half-wave resonances ($M = 3$, $M = 4$) are resonant at frequencies that are very close to the specified window frequencies. For example, if a free spectral range of 35 GHz is assumed, the third order resonance is at 105 GHz and the fourth is at 140 GHz. Since the precise window frequencies are not critical, the 105 GHz frequency would serve the purpose as well as 104 GHz. Further, with a FSR of only 35 GHz, the plate spacing is about 4.2 mm, and the walkoff geometric losses can be kept to 0.12 dB at an etalon tilt angle of $\theta = 15^\circ$.

The reflectivity of the etalon need be only 0.7, which gives a midband rejection at 118 GHz of 15 dB. In terms of reflection loss for the 118 GHz band, this amounts to 0.14 dB.

Design parameters are summarized in Table 1. An aperture of 70 mm is assumed for the 104 GHz band. It should be noted that walkoff diffraction losses are negligible because the plate spacing is only 4.2 mm.

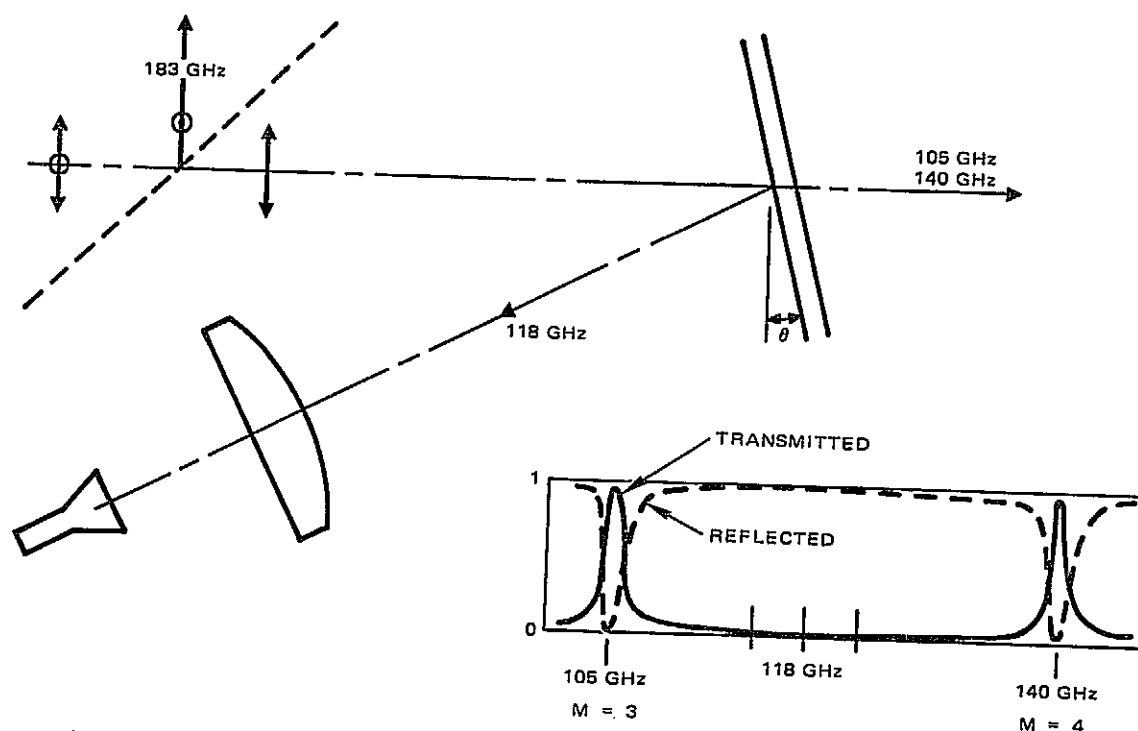
References

1. P. F. Goldsmith, "Quasi-Optical Feed System for Millimeter Wave Radiometer, "BSTJ, Volume 56, No. 8, pp. 1483 - 1502, October 1977.
2. P. G. Wannier, J. A. Arnaud, F. A. Pelow, and A. A. M. Saleh, "Quasi-Optical Band-Rejection Filter at 100 GHz", Review of Scientific Instruments, Vol. 47, No. 1, January 1976, pp. 56-59.
3. A. A. M. Saleh, "An Adjustable Quasi-Optical Bandpass Filter - Part I, Theory and Design Formulation, " MTT Transactions, Vol. 22, No. 7, July 1974, pp. 728-734.
4. A. A. M. Saleh, "An Adjustable Quasi-Optical Bandpass Filter - Part II, Practical Considerations, " MTT Transactions, Vol. 22, No. 7, July 1974, pp. 734-739.

TABLE 1. FABRY-PEROT MULTIPLEXER FILTER CHARACTERISTICS

Type	Fabry-Perot parallel plate
Mode	Reflect 118 GHz, transmit 105, and 140, or 104.6 (M = 3), 139.5 (M = 4)
3 dB bandwidth	4 GHz
Free spectral range	35 GHz
Beam waist	22.5 mm
Aperture diameter	70 mm
Walkoff diffraction loss	0 dB
Reflectivity	0.7
Plate spacing	4.2 mm
Ohmic loss at 105, 140 GHz	0.1 dB
Reflective loss at 114 to 122 GHz	0.14 dB (118 GHz feedthrough = -15 dB)
Walkoff geometric loss	0.12 dB at $\theta = 15^\circ$ tilt

70532-65T



70532-64

FIGURE 1. FABRY-PEROT MULTIPLEXER FILTER FOR 104, 118, AND 140 GHz BANDS

4. Receiver Design

4.6 Quasi-optical Diplexers

4.6.3 SEPARATION OF 118 GHz BAND FROM WINDOW BAND USING RESONANT GRID DICHROIC FILTERS

Single and double-sheet resonant grid diplexers can be used to separate the window bands from the 118 GHz band. These devices have the advantage of operating well at 45° incidence.

In its simplest form, a dichroic filter is a thin sheet of conducting material perforated with slots that are resonant in the desired frequency band. Two or more sheets resonant at the same frequency or at separate frequencies may also be used to produce sharper passband characteristics. Excellent design approximations can be obtained from a mathematical model based on a planar array of identical waveguide elements. In the present application, the resonant grid will be placed at an angle of 45° simulating a planar array scan angle of 45°. Given the slot dimensions and their periodicity in the array environment, the computer program determines the transmission and reflection coefficients of the assumed modes at the aperture. The transmission and reflection coefficients are determined by a mode matching technique invoking the continuity of the tangential fields of the waveguide modes and the aperture fields expressed in terms of "free space" modes.^(1,2) The slot configuration and array geometry are given in Figure 1. At the slot resonant frequency, incident energy is transmitted without loss. On either side of resonance the slots become reactive and the complete sheet becomes reflective.

The transmission and reflection characteristics of a resonant grid backed by a 0.002 inch thick material of dielectric constant $\epsilon_r = 3.1$ (Kapton) are shown in Figure 2. Clearly, the resonant grid passes the 135 to 145 GHz frequency band and rejects all other frequencies accordingly. The final dimensions of the resonant grid are arrived at after several iterations; in each case care must be taken that the selected slot spacings do not give rise to grating lobes in the region of interest. Grating lobe formation is avoided by selecting $DX < 2\lambda/(1 + \sin\theta)$ where λ is the wavelength corresponding to the highest frequency of the band and θ is the maximum scan angle of the array.

The design of the resonant grid for the 110 to 125 GHz passband is much more difficult since the rejection bands (102 to 106 GHz and 135 to 145 GHz) are relatively close to the passband. Single sheet and double designs for various combinations of slot dimensions and two different slot spacings are shown in Figure 3.

The steepness of the slope between the transmission and rejection bands may be increased by placing a second sheet of identical design a suitable distance behind the first sheet. Since the resonant frequency is the same for the two sheets, all the energy will be transmitted irrespective of the spacing between the two sheets at the resonant frequency. Variation of the spacing between the sheets changes the effective bandwidth of the passband. This is illustrated in Figure 3. The curves clearly show that the bandwidth may be increased by increasing the spacing between the two sheets, while the slope between the transmission and rejection bands remains essentially unchanged. The slope may be made steeper, however, by the addition of identical sheets.

References

1. N. Amitey, V. Galindo, and C. P. Wu, Theory and Analysis of Phased Array Antennas, Wiley, New York, 1972.
2. S. W. Lee, W. R. Jones, and J. J. Campbell, "Convergence of Numerical Solutions of Iris-Type Discontinuity Problems," IEEE Transactions MTT, Vol. MTT-19, No. 6, June 1971, pp. 528, 537.
3. A. A. M. Saley, R. A. Semplak, IEEE Transactions on Antennas and Propagation, Vol. AP-24, No. 6, November 1976.
4. J. A. Arnaud and F. A. Pelow, "Resonant Grid Quasi-Optical Diplexers," BSTJ, Vol. 54, No. 2, February 1975.

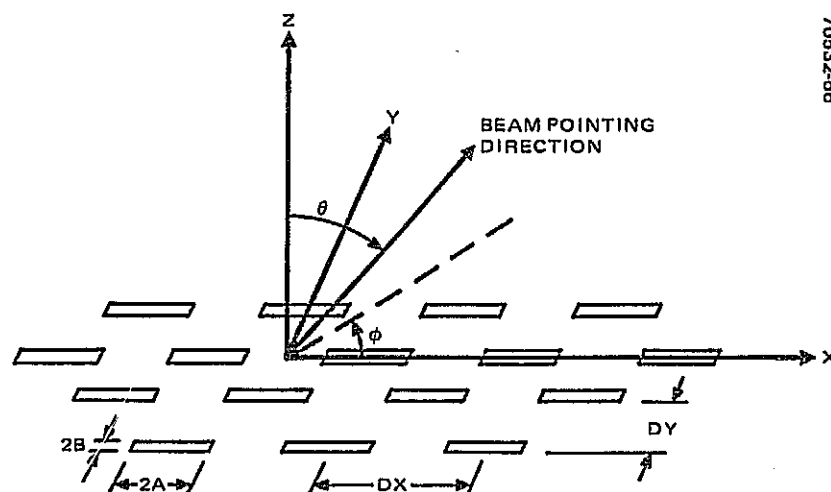


FIGURE 1. SLOT CONFIGURATION AND ARRAY GEOMETRY

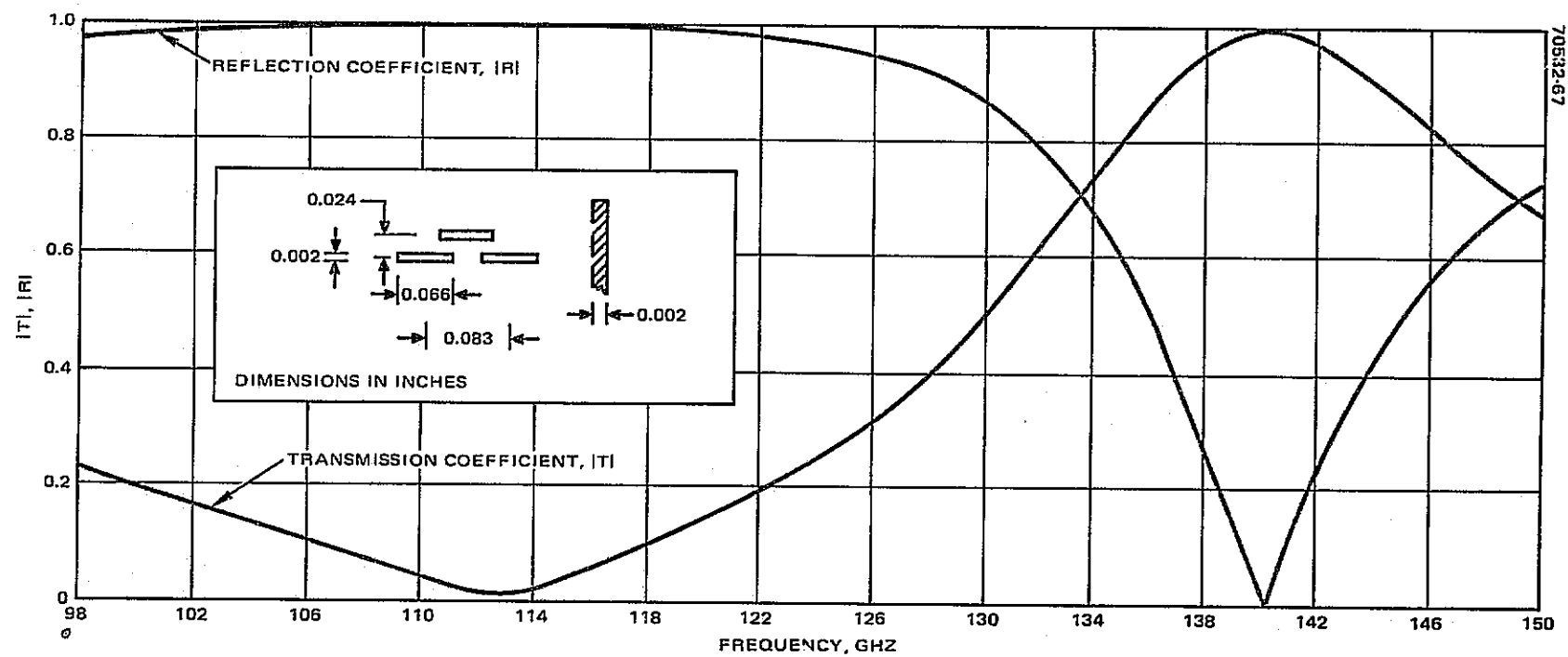


FIGURE 2. 104/140 GHz DICHROIC FILTER RESPONSE

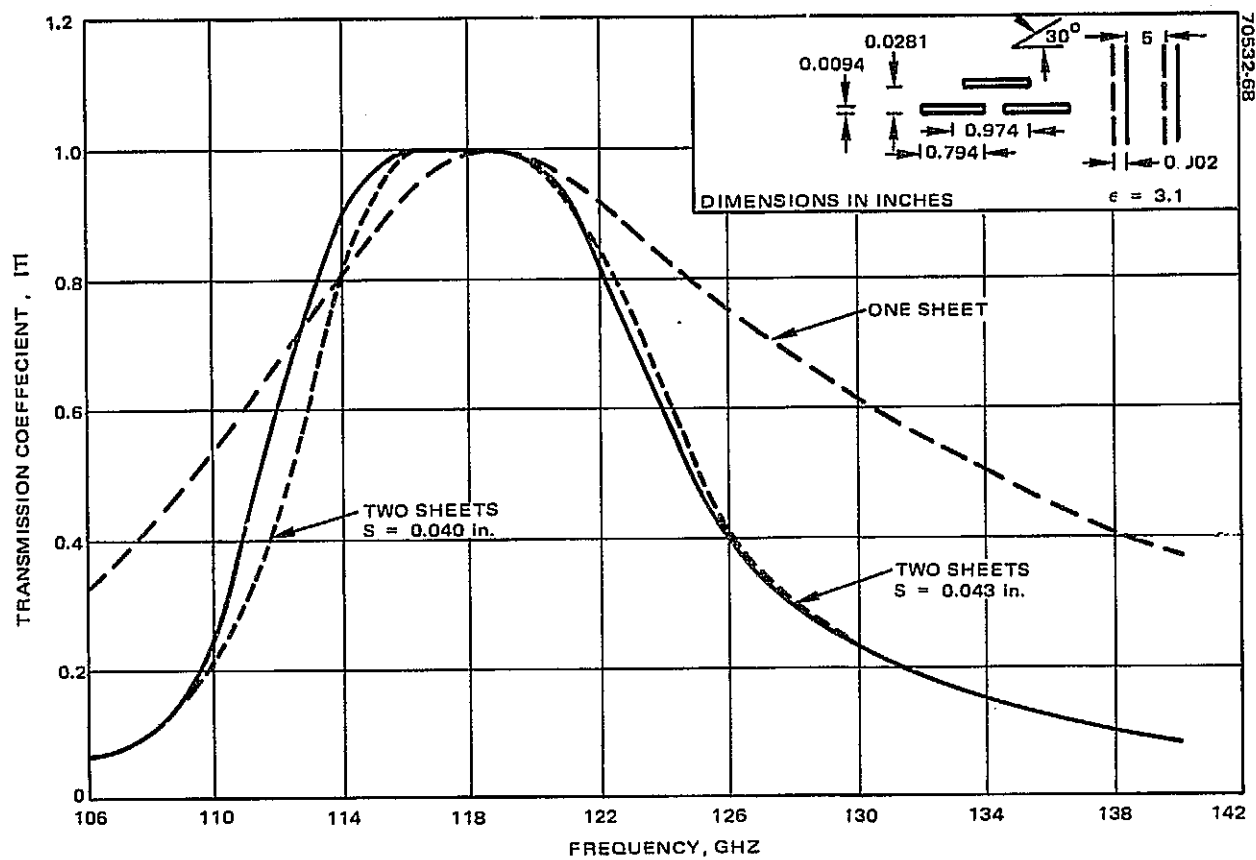


FIGURE 3. 118 GHz DICHROIC FILTER RESPONSE

5. ANTENNA INTERFACE

5.1	ILLUMINATION PATTERNS	5-4
5.2	MASR ANTENNA AND FEED GEOMETRY	5-6
5.3	QUASI-OPTICAL QUADRA PLEXING	5-10
5.4	ANTENNA CALIBRATION	5-12
5.5	SYSTEMATIC ERRORS	5-20

5. ANTENNA INTERFACE

Interfaces between the radiometer receiver and antenna consist of mechanical/thermal, electrical, feed, and calibration. Special requirements of the feed interface include illumination with a quasi-optical quadraplexer.

One of the primary tasks of MASR system engineering is to establish interface specifications between the antenna and radiometer receiver subsystems. Such interfaces as mechanical/thermal and electrical, including bus power and telemetry and command, are routinely achieved during system definition phases of flight programs. Establishing those interfaces at this time in the concept phase is premature. However, for the present study two critical areas require preliminary interfaces to be established: the feed interface and the antenna calibration interface.

Feed Interface. The concept of using beam optics design formalism in the receiver subsystem has assumed the use of Gaussian illumination to facilitate design procedures and to minimize diffraction losses. This concept implies that illumination patterns are generated in the receiver subsystem and are the responsibility of the receiver subsystem engineering. Thus, contrary to a conventional antenna interface such as a waveguide flange, the MASR antenna interface consists of a set of four complex beam parameters at the focus of the antenna.

Corrugated horns with quarter-wave spacings carrying the HE_{11} mode produce nearly symmetric circular Gaussian field amplitude patterns. Gaussian design formalism is therefore appropriate in quasi-optical feeds for MASR. Separate horns may optimally generate the pattern for each frequency band, and these bands can be combined with a quasi-optical quadraplexer. Topics 5.1, 5.2, and 5.3 address the details of the feed interface.

Antenna Calibration. The radiometer receiver measures antenna temperature continuously with a ΔT resolution of typically 0.2° and an absolute temperature accuracy of $\pm 1^\circ K$. Antenna characteristics must be known to relate the radio-metrically measured antenna temperature to the desired main beam radiation temperature. Briefly, the calibration procedure consists of three steps. A "cold" look into space will yield a residual antenna temperature caused by ohmic losses. A relative antenna pattern measurement will allow observations to be corrected for sidelobe introduced errors. Finally, comparison between radiosonde and MASR temperature measurements will yield an accurate absolute calibration. Topic 5.4 addresses the details of the antenna interface.

~~RECHING~~ PAGE BLANK NOT FILMED

5. Antenna Interface

5.1 ILLUMINATION PATTERNS

Corrugated horns with quarter-wave spacings carrying the HE_{11} mode produce nearly symmetric circular Gaussian field amplitude patterns. Gaussian optical design formalism is therefore appropriate in quasi-optical feeds for the MASR antenna.

By convention, an antenna pattern of a receiving antenna is defined by reciprocity to be identical to the pattern generated by a transmitter having the identical antenna configuration. The MASR receiver antenna pattern is thus a function of the illumination pattern of the feed horn.

For an antenna of infinite extent, a Gaussian illumination profile produces a far field pattern which has no sidelobe structure. The main beam field intensity decreases monotonically off axis in a characteristic Gaussian pattern. In practice, a Gaussian illumination pattern is truncated by the antenna edge which introduces sidelobes in the far field. Illumination patterns generated by computers through iterative processes to obtain minimum sidelobes are very nearly Gaussian profiles.

The radiation from a horn can be made polarization symmetric through the use of corrugations spaced $\lambda/4$ apart near the radiating aperture.⁽¹⁾ The field pattern of the HE_{11} mode in the region of the horn is then given by^(2, 3)

$$g(r) = E_0 J_0 \left[\frac{2.405 r}{a} \right]$$

(r and a are defined in Figure 1).

The Gaussian approximation of this function is

$$g(r) = E_0 \exp(-r^2/w_0^2)$$

where w_0 is the beam waist when the field is e^{-1} times the field amplitude on axis and r is the variable distance from the z -axis. If we let these two functions be equal at the e^{-1} point, we have the relations $r = w_0$

and

$$2.405 \frac{w_0}{a} = 1.7538$$

$$w_0 = 0.73a$$

Thus,

$$g(r) = E_0 \exp(-r^2/0.53a^2)$$

Both the Bessel and Gaussian functions are plotted in the Figure 1. It can be seen that very close agreement is achieved by the proper choice of Gaussian profile. The tail of the Gaussian is truncated at $r = a$ whereas the Bessel function goes to zero at the walls. This inconsistency is the greatest anomaly in the Gaussian

approximation of the field in the region of the horn. However, in the far field of the horn, for large values of z , very close approximations to Gaussian have been demonstrated. (4) Side lobe levels of -40 dB and antenna beam efficiencies of 98 percent are common using these methods. The close approximation of actual optimized fields in high performance antennas to Gaussian fields make the use of Gaussian optical design formalism not only warranted but actually preferred to other design approaches.

References

1. C. M. Knop and H. J. Wiesenfarth, "On the Radiation from an Open-Ended Corrugated Pipe Carrying the HE_{11} Mode," IEEE Transactions, PGAP, AP-20, September 1972, pp. 644-648.
2. P. S. Neelakantaswamy, K. K. Gupta, and D. K. Banerjee, "A Gaussian-Beam Launcher for Microwave Exposure Studies," IEEE Transactions, PGMTT, MTT-25, May 1977, pp. 426-428.
3. C. Aubry and D. Bitter, "Radiation Pattern of a Corrugated Conical Horn in Terms of Laguerre-Gaussian Functions," Electronics Letters, Vol. 11, No. 7, April 1975, pp. 154-156.
4. C. Dragone, "An Improved Antenna for Microwave Radio Systems Consisting of Two Cylindrical Reflectors and Corrugated Horn," BSTJ, Vol. 53, No. 7, September 1974, pp. 1351-1377.

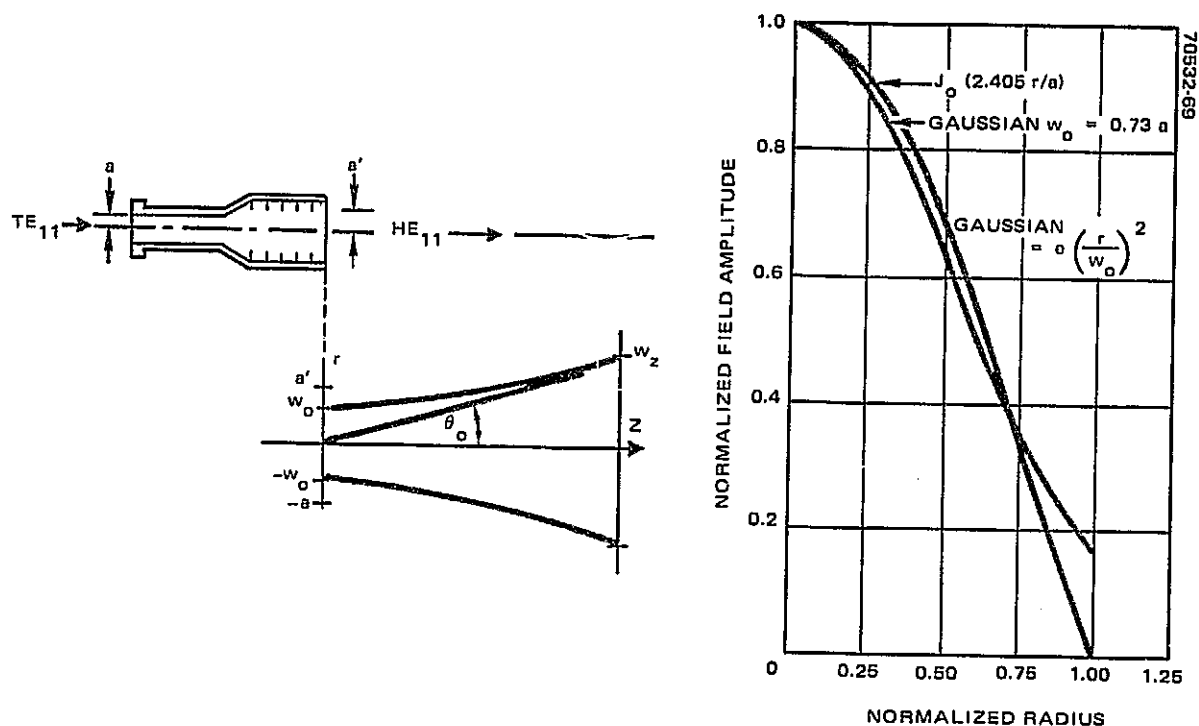


FIGURE 1. SCALAR HORN ILLUMINATION PATTERN COMPARED TO GAUSSIAN

5. Antenna Interface

5.2 MASR ANTENNA AND FEED GEOMETRY

The recommended antenna configuration for MASR is a symmetrical cassegrain reflector system. The main reflector truncates the Gaussian illumination at the 99 percent beam energy points, and forward spillover is limited to -32 dB down from the total beam energy.

A detailed tradeoff study for the MASR antenna system has determined that the symmetrical cassegrain antenna is lighter, more compact, and has a better pattern for given surface figure than the offset paraboloidal reflector.⁽¹⁾ The basic antenna layout is shown in Figure 1. The primary reflector is a 4.4 meter diameter parabolic reflector which has a prime focus at 1.54 meter from the center of the dish. The secondary hyperboloidal reflector is disposed between the primary reflector and focus such that an external focus is formed 0.16 meter behind the surface of the antenna. The external focus is the antenna feed point for the four radiometer frequency bands, 104 GHz being the lowest and 183 GHz being the highest frequency.

Since the antenna feed point is located 1.64 meter from the center of the subreflector, the energy distribution characteristics at the external focus will be determined by diffraction.

A special property of Gaussian beams is that the relative intensity measured at a point is equal to the fractional power for all radii beyond that point. In other words, the relative intensity at a point at the edge of the subreflector is equal to the fractional spillover. In the presently recommended design, the illumination of the primary is controlled by a maximum subreflector diameter of 0.198 meter, while the spillover is limited by a total subreflector diameter of 0.25 meter. Thus, the relative intensity at 0.198 meter is

$$\frac{I}{I_0} = \exp\left(\frac{-2r^2}{w_z^2}\right) = 0.01$$

where

$$r = \frac{0.198}{2}$$

Solving for w_z we have

$$w_z = 0.065 \text{ m}$$

Having determined that the e^{-1} illumination circle of points is 6.5 cm from the center of the center of the subreflector, the spillover or energy which falls beyond the 0.25 meter outer diameter is

$$\frac{P}{P_0} = \exp\left[\frac{-2(0.125)^2}{w_z^2}\right]$$

or about -32 dB down from the total power. It has been determined that spillover of these levels does not seriously affect the radiometer performance (Topic 5.3).

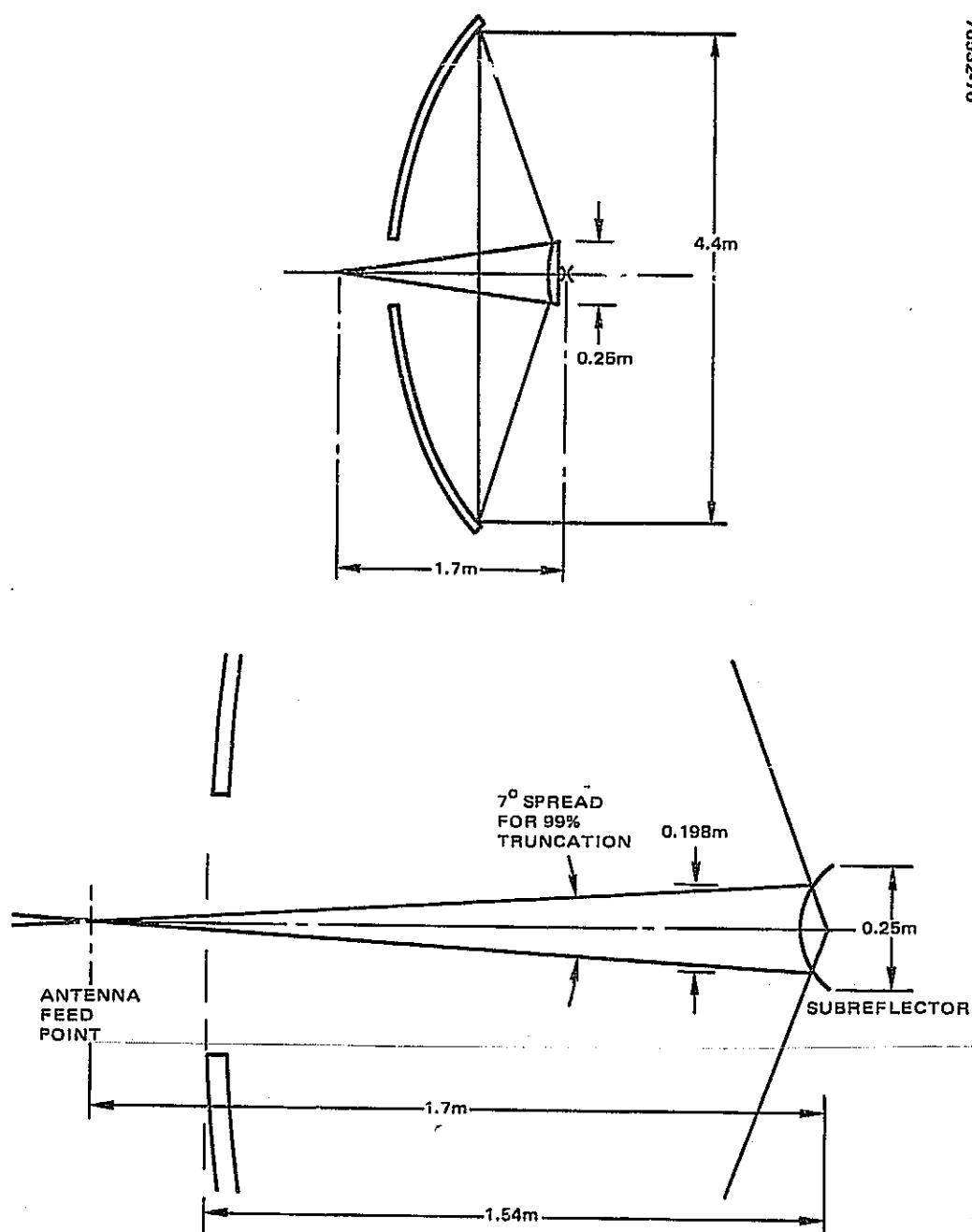


FIGURE 1. MASR ANTENNA LAYOUT

The external focus is the feed point of the cassegrain antenna. The feed point consists of four coaxial beams representing four radiometric bands. The 104 GHz band has a beam diameter of 63.4 mm to the 1 percent intensity points, and the 183 GHz band has a beam diameter of 39.5 mm. These beam diameters are determined by 1) the illumination requirement of the subreflector and 2) the laws of diffraction. The illumination requirement is shown in Figure 2. Note that the "amplitude" profile is shown truncated at 0.1 at the beam limits. The intensity at these limits is 1 percent, and 99 percent of the beam energy is contained within these limits. Also note that the illumination pattern requirements holds for all frequencies. Thus, the diffraction limited "focus" can be determined from the Gaussian formalism described in Topic 3.7. The fixed conditions are the values of w_z and the distance F from the subreflector to the focus which fixes the diffraction half angle

$$\theta_{1/2} = \frac{w_z}{F} = 0.040 \text{ rad}$$

From Topic 3.7 we have the relation

$$\theta_{1/2} = \frac{\lambda}{\pi w_0} \text{ or } w_0 = \frac{\lambda}{\pi \theta}$$

from which the beam waists for each wavelength can be determined. The values of beam diameters at the 99 percent beam energy points are obtained by multiplying the beam waist values by 3.035.

Figure 3 illustrates how the four frequency bands give different beam diameters at the antenna focus. Table 1 lists the values of beam waist, beam diameters, and confocal parameters for each of the four frequency bands. The confocal parameter values are essential in designing the quasi-optical quadraplexer which is described in Topic 5.3.

Reference

1. A. T. Villeneuve, "Microwave Atmospheric Sounding Radiometer Antenna Feasibility Study," Final Report, Contract NAS5-24087, December 1977.

TABLE 1. BEAM PARAMETERS FOR FOUR RADIOMETER BANDS

Frequency, f, GHz	Wavelength, λ , mm	Half Angle, $\theta_{1/2}$, rad	Beam Waist, w_0 , mm	Beam Diameter at 1% Intensity, mm	Confocal Parameter, b, mm
183	1.64	0.040	13.05	39.5	326
140	2.14	0.040	17.03	51.6	425
118	2.54	0.040	20.2	61.2	505
104	2.88	0.040	22.9	63.4	572

70532-73T

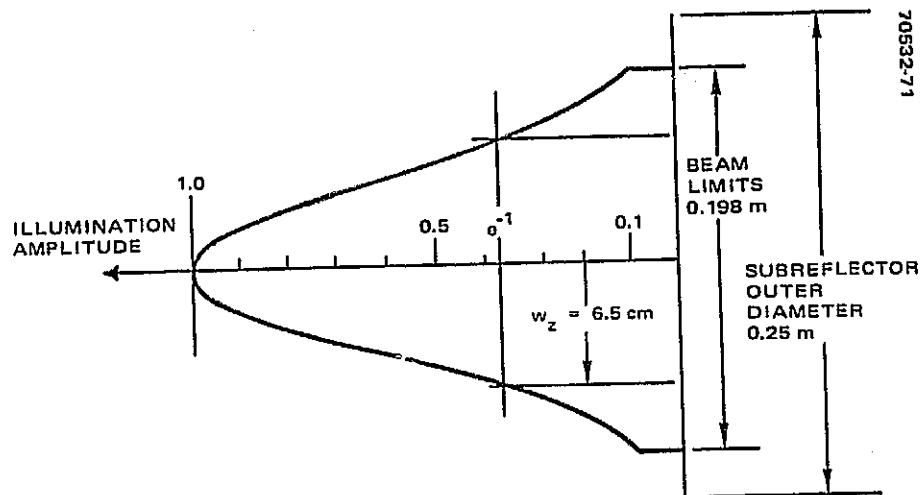


FIGURE 2. ILLUMINATION PROFILE REQUIREMENT FOR ALL FREQUENCIES

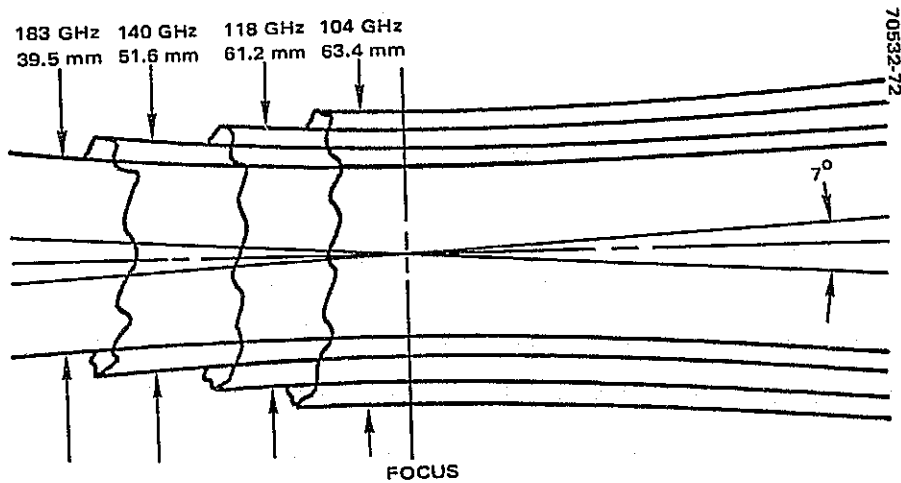


FIGURE 3. BEAM DIAMETERS AT WAIST AT EXTERNAL FOCUS
FOR 99 PERCENT TRUNCATION

5. Antenna Interface

5.3 QUASI-OPTICAL QUADRAPLEXING

Four separate feed horns for four radiometric bands can be imaged to common feed point through quasi-optical techniques. Feed losses for the 183 GHz band can be kept below 0.5 dB and below 1 dB for the worst band.

The four band coaxial beam geometry, illustrated in Topic 5.2, enters the radiometer and must be spectrally separated and directed to the four receivers. The techniques of quasi-optics were selected for this design example since losses can be minimized. The demultiplexing technique is to use spectrally sensitive mirrors to create images of the antenna focus suitable for the individual bands. Figure 1 shows that mirrors imposed in front of the real antenna focus, or feed point, to reflect images to the 183 GHz and 118 GHz feed horns. Lens images or relay images are formed for the 104 and 140 GHz bands.

The first mirror in the system was determined to be a wire grid polarizer for good reasons: it is the lowest loss quasi-optical element which is capable of separating two beams (Topic 4.6); and it takes advantage of the energy in both polarizations which would otherwise be lost in a more conventional diplexer. The second mirror may be either a Fabry-Perot filter or a resonant grid filter. Also, the third mirror may be either a Fabry-Perot or resonant grid device.

Estimated feed losses for each band are listed in Table 1. Losses for the 183 GHz band assume the use of a quasi-optical balanced mixer which also uses a wire grid. A horn loss of 0.25 dB for each band is assumed. This loss is based on measurements of the throughput loss of two identical horns operating against a parabolic mirror at conjugate foci. The total loss of two horns at 140 GHz was 0.5 dB.

TABLE 1. ESTIMATED FEED LOSSES FOR FOUR RADIOMETER BANDS

Band	Wire Grid 1	Wire Grid*	Fabry-Perot 2	Fabry-Perot 3	Resonant Grid 2	Resonant Grid 3	Horn Loss, dB	Total Loss, dB
183	<0.1	<0.1					0.25	<0.45
140	<0.1		0.1**	0.1**			0.25	<0.55
	<0.1				0.5	0.1**	0.25	<0.95
118	<0.1		0.14				0.25	<0.5
	<0.1				0.5		0.25	<0.85
104	<0.1		0.1*				0.25	<0.55
	<0.1				0.5	0.1**	0.25	<0.95

*Assumed using quasi-optical balanced mixer.

**Midband loss, ADD 1 dB at band edge.

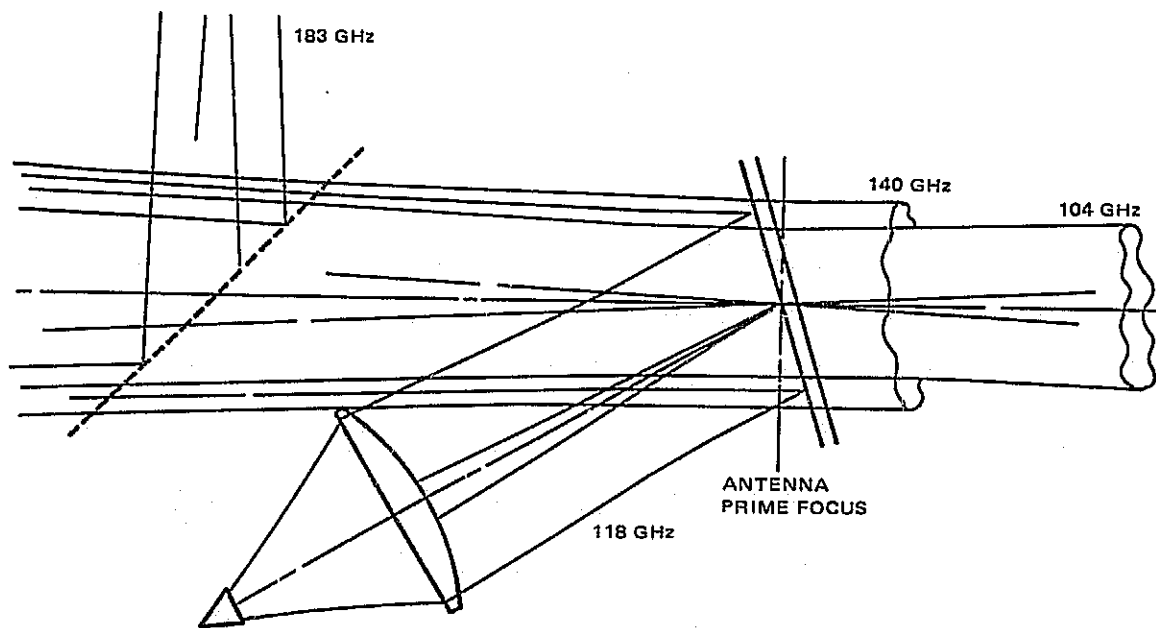
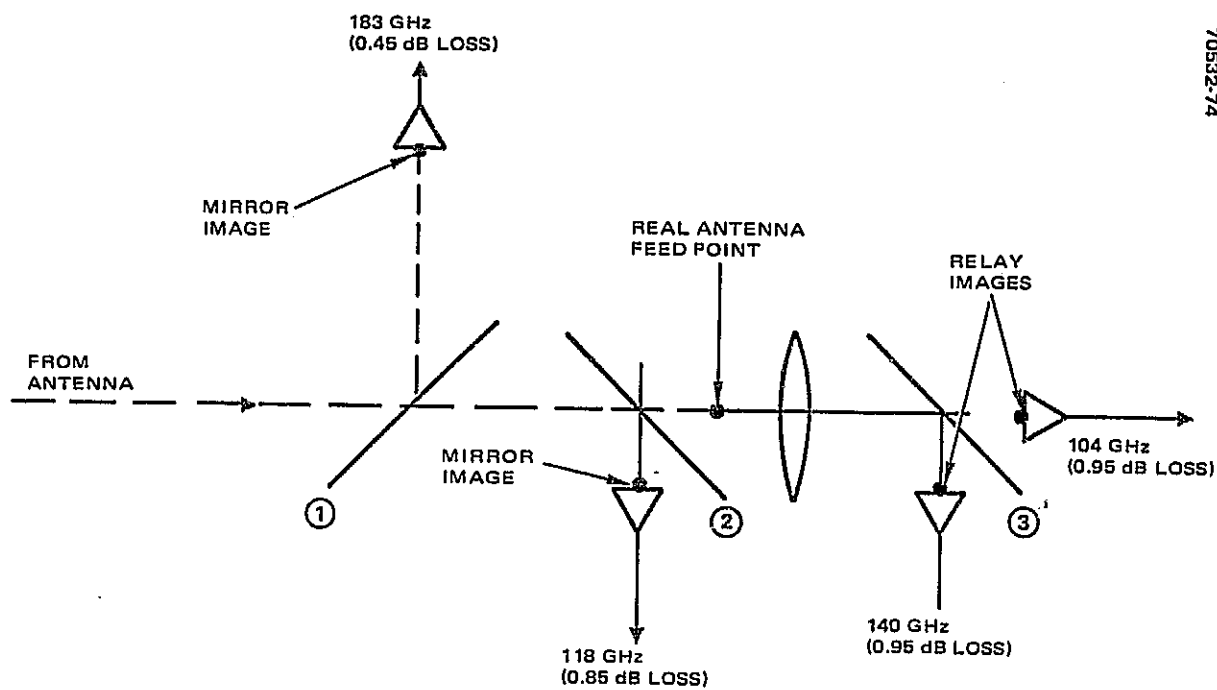


FIGURE 1. QUASI-OPTICAL FEED SCHEMATIC

5. Antenna Interface

5.4 ANTENNA CALIBRATION

Antenna performance must be known to relate the radiometrically measured antenna temperature to the desired main beam radiation temperature.

Antenna Temperature. The measured antenna temperature at the radiometer calibration reference plane for each radiometric channel is related to the radiation brightness temperature in all directions by the equation

$$T_A = \frac{k_r}{4\pi} \iint D(\theta, \phi) T_B(\theta, \phi) d\Omega + (1 - k_r) \bar{T}_O \quad (1)$$

where

T_A = antenna temperature

k_r = antenna ohmic efficiency

D = antenna directivity

T_B = radiation brightness temperature

(θ, ϕ) = polar and azimuthal angles about the antenna

$d\Omega$ = element of solid angle

$\bar{T}_O$ = average physical temperature of the antenna reflecting surfaces, weighted by the feed system illumination pattern

The antenna directivity is defined as the ratio of the power intensity produced by the antenna in the given direction to the mean value of the intensity in all directions.

The first term of equation 1 is the contribution to the antenna temperature from energy received by the antenna, reduced by the ohmic efficiency. The second term is due to radiation emitted by losses present in the antenna.

There are two sources of antenna resistive loss. The first is due to small losses present in the antenna reflecting surfaces, and the second is due to the portion of subreflector blockage energy which is terminated in the mixer. Both of these losses contribute to k_r .

An alternate representation of this resistive loss due to blockage is obtained by considering the mixer diode as a thermal radiator at a temperature determined by diode characteristics and operating point. Some of the thermal energy emitted by the mixer will be reflected back into the feed by the subreflector, raising the antenna temperature (see section 5.5).

$D(\theta, \phi)$ is composed of a main beam, near sidelobes, and remote sidelobes. Near sidelobes are defined as those which fall on the disc of the earth ($T_B \approx 250^\circ\text{K}$) and

remote sidelobes are directed to cold space ($T_B \approx 2.5^\circ\text{K}$). The antenna temperature can now be expressed as

$$T_A = \bar{T}_m k_r \eta_m + \bar{T}_{ns} k_r \eta_{ns} + \bar{T}_{rs} k_r \eta_{rs} + (1 - k_r) \bar{T}_o \quad (2)$$

where

η_m = portion of the pattern energy in the main beam

$$= \frac{1}{4\pi} \iint_{\text{main beam}} D(\theta, \phi) d\Omega$$

and where

$\bar{T}_m$ = weighted average main beam radiation temperature

$$= \frac{1}{4\pi\eta_m} \cdot \iint_{\text{main beam}} D(\theta, \phi) T_B(\theta, \phi) d\Omega$$

similarly for $\bar{T}_{ns}$, η_{ns} and $\bar{T}_{rs}$, η_{rs} , the near and remote sidelobe temperatures and percent energy contributions,

$$\eta_m + \eta_{ns} + \eta_{rs} = 1.$$

The first term in Equation 2 is the contribution to the antenna temperature due to the brightness temperature in the main beam. The second term represents energy collected by antenna near sidelobes from the earth. The third term accounts for far sidelobe energy from cold space and the fourth term is due to energy emitted by the antenna reflecting surfaces. The energy collected from the sun in the antenna remote sidelobes is negligible.

To accurately determine the main beam brightness temperature T_m from T_A , other factors affecting antenna temperature must be known. This requires knowledge of the antenna performance characteristics.

Calibration Procedure. Range measurement of the MASR antenna is impractical due to the large range length required to satisfy the far field criterion.

$$R \geq \frac{2d^2}{\lambda}$$

where R is the range length, d is the aperture diameter, and λ is the wavelength. For the MASR antenna at 183 GHz, the minimum range length is about 15 miles. Ground reflections and atmospheric effects would degrade measurement accuracy. However, the antenna performance can be measured after launch.

Briefly, the calibration procedure consists of: 1) a "cold look" into space with the main antenna system; 2) an approximate antenna pattern measurement; and 3) an atmospheric temperature measurement by both radiosonde and MASR.

The cold look is used to determine the antenna ohmic efficiency k_r . The pattern measurement yields an estimate of the ratio of η_m to η_{ns} and an energy distribution in the near sidelobes, which is used to correct MASR measurements for sidelobe contributions in a relative manner. Finally, comparison between MASR and radiosonde measurements enables the total amount of pattern energy on the disc of the earth ($\eta_m + \eta_{ns}$) to be accurately determined. The calibration is complete.

1) Cold Look. With the antenna looking at cold space, the first term in Equation 1 is near zero, which exposes the antenna resistive loss radiation $(1-k_r)\bar{T}_0$. To minimize collected energy the earth, moon, and sun should be located behind the antenna. This posture could be accomplished by positioning the spacecraft looking away from the earth. Reflecting surface temperatures needed to calculate $\bar{T}_0$ can be obtained from an antenna thermal model, allowing the determination of k_r . (6)

2) Pattern Measurement. the antenna pattern is needed over an angle somewhat larger than the disc of the earth to allow the calculation of energy collected by antenna sidelobes from the earth during observations. The overall system temperature accuracy is affected very little by the accuracy of the measured antenna pattern, and an approximate pattern is all that is required. The pattern can be measured in orbit using either a ground-based transmitter or astronomical sources.

A ground transmitter measurement is simple in concept and has the capability of wide dynamic range with accurate control of transmitted power. Atmospheric absorption in the oxygen and water vapor high frequency IF channels is less than 10 dB and should not be a problem. The transmitted signal can be modulated to improve sensitivity against the thermal background of the earth.

However, the demands placed on the antenna positioning system by a transmitter measurement are severe. To map the main beam using a transmitter would require an antenna positioning accuracy of about one-tenth beam width or 0.05 mr at 183 GHz. The presently planned 0.15 mr would not be sufficient. Also, due to the fine structure of the antenna pattern, a large number of data points would be required.

Pattern measurement using the moon and the sun as millimeter wave sources eliminates the disadvantages of a ground-based transmitter measurement while yielding results accurate enough for our use. An averaged pattern, rather than a point-by-point mapping, can be obtained by deconvolving scans across the moon with antenna patterns containing various amounts and distributions of sidelobe energy. Constraints on a correctly deconvolved scan are that no energy come from beyond the disc of the moon and that there be no limb brightening at the dark edge. Average sidelobe levels to 45 dB below the main beam can be mapped using the moon. The sun can be used to extend this measurement to -60 dB.

This procedure is described in Reference 1 where it was used to measure the pattern of the National Radio Astronomy Observatory 11 meter millimeter-wave telescope operating at 214 GHz. The pattern information obtained was used to correct solar maps for sidelobe introduced errors.

The orbital motion of the satellite can be used to scan the antenna beam across the moon, eliminating the problems with antenna positioning inherent in a ground

transmitter measurement. The orbital angular rate is 0.07 mr/sec, or about 7 seconds/beamwidth at 183 GHz, which is a convenient rate.

Pattern measurement using large sources such as the moon and the sun does not yield information about the fine structure of the antenna pattern, but an averaged pattern is all that is required to make an estimate of the temperature contribution of the sidelobe energy during earth observations.

3) Radio-sonde Comparison. The amount of the total antenna energy which falls on the earth ($\eta_m + \eta_{ns}$) must be known, and can be determined by comparing radiometric and direct atmospheric measurements. The ratio of η_m to η_{ns} is known from the pattern measurements. This ratio, along with observation over the disc of the earth, allows the radiometric observations in the area of the radio-sonde measurement to be corrected for relative sidelobe contributions. At this point, the estimate of ($\eta_m + \eta_{ns}$) is adjusted to bring the radiometric measurements into agreement with the radio-sonde measurements. This procedure assumes that $\bar{T}_m \doteq \bar{T}_{ns}$, which is accurate.

The amount of energy on the earth outside the main beam will be a function of beam position on the earth. The antenna patterns are used to calculate changes in η_{ns} and η_{rs} for various beam positions.

Temperature Accuracy Estimate. To calculate an estimate of MASR system temperature accuracy, nominal values of antenna performance and projected temperatures for the worst case at 183 GHz will be used to determine the accuracy of the calibration procedure previously discussed. Then the uncertainty in system performance parameters obtained during calibration will be used to calculate overall system temperature accuracy.

1) Antenna Performance Estimate (183 GHz)

	<u>Power Breakdown, %</u>	<u>Temperature</u>
Resistive losses	1	antenna $\bar{T}_o$
Central blockage, 2%	1	earth
	1	antenna $\bar{T}_o$
Spar blockage	2	space
Spillover	1	earth
Random scattering, 5%	4	earth
	1	space
Sidelobes, 4%	2	earth
	2	space
Main beam	85	earth

Performance Parameters (nominal values)

$$k_r = 0.98 \text{ (from resistive loss and central blockage)}$$

$$\eta_m = 0.8674$$

$$\eta_{ns} = 0.0816$$

$$\eta_{rs} = 0.0510$$

2) Cold Look. To estimate the error in the measurement of $(1 - k_r)$, Equation 2 is solved for the term

$$(1 - k_r) = \frac{1}{\bar{T}_O} \left[T_A - (\bar{T}_m k_r \eta_m + \bar{T}_{ns} k_r \eta_{ns} + \bar{T}_{rs} k_r \eta_{rs}) \right]$$

Since

$$\bar{T}_m = \bar{T}_{ns} = \bar{T}_{rs} = \bar{T}_{\text{signal}} \ll \bar{T}_O$$

$$(1 - k_r) = \frac{1}{\bar{T}_O} \left[T_A - (T_{\text{signal}}) \right]$$

Using nominal values and estimated errors of $\bar{T}_O$, T_A , and T_{signal} , the 1σ error in $(1 - k_r)$ is the rss (root sum of the squares) of the errors due to each component.

Error	$\Delta(1 - k_r)$
$\bar{T}_O = 200 \pm 20^\circ\text{K}$	0.002
$T_A + \Delta T_A = T_A \pm 0.7^\circ\text{K}$	0.0035
$T_s + \Delta T_s = T_{\text{signal}} \pm 0.4^\circ\text{K}$	0.002
$(1 - k_r) = 0.02 \pm 0.0045$	

To correctly compute the expected value for T_A , the radiometric temperature scale, not the depressed radiometric temperature scale, should be used to add the contributions from $\bar{T}_O$ and T_{signal} (Topic 3.5). At 183 GHz, cold space emits very little radiation, and T_A is comprised largely of antenna resistive loss radiation.

3) Pattern Measurement. It is anticipated that the pattern measurement using the moon and the sun will be accurate enough to determine the ratio η_m to η_{ns} to within 20 percent.

$$\frac{\eta_m}{\eta_{ns}} = 10.6 \pm 2$$

4) Radioonde Measurement. Using the radioonde data and the atmospheric model, the main beam radiation temperature $\bar{T}_m$ can be predicted to within 1°K. This measurement will be conducted in an area of slowly varying temperature, $\bar{T}_m - \bar{T}_{ns} \leq 10^\circ\text{K}$. By making observations around the test point, $\bar{T}_{ns}$ can be determined to within 5°K.

The purpose of the joint MASR/radioonde observation is to determine the sum $\eta_m + \eta_{ns}$, the percentage of the pattern energy falling on the disc of the earth. To solve Equation for $\eta_m + \eta_{ns}$, the values of $\bar{T}_m$ and $\bar{T}_{ns}$ are combined into an average earth temperature, $\bar{T}_e$, known to 1°K due to the small effect of the uncertainty in $\bar{T}_{ns}$.

$$\eta_m + \eta_{ns} = \frac{T_A - \bar{T}_{rs} k_r \eta_{rs} - (1 - k_r) \bar{T}_o}{\bar{T}_e k_r}$$

Summing the errors as before

Error	$\Delta(\eta_m + \eta_{ns})$
$T_A = 245.5 \pm 0.7^\circ\text{K}$	0.0036
$\bar{T}_{rs} = 4.5^\circ \pm 0.5^\circ\text{K}$	0.0001
$\bar{T}_o = 200^\circ \pm 20^\circ\text{K}$	0.0020
$\bar{T}_e = 259.5 \pm 1^\circ\text{K}$	0.0037
$k_r = 0.98 \pm 0.0045$	0.0008
$\eta_m + \eta_{ns} = 0.949 \pm 0.0056$	

The calibration is complete.

5) Overall System Accuracy. To determine the uncertainty in $\bar{T}_m$ during observations, Equation 2 is solved for this quantity.

$$\bar{T}_m = \frac{1}{k_r \eta_m} [T_A - T_{ns} k_r \eta_{ns} - T_{rs} k_r \eta_{rs} - T_o (1 - k_r)]$$

Nominal values and uncertainties are known for all terms on the right hand side, and the error in $\bar{T}_m$ can be found as before. However, because the sum $\eta_m + \eta_{ns} + \eta_{rs} = 1$ without error, there are only two independent errors involving the three η , which are the error in the ratio η_m to η_{ns} and the error in the sum $\eta_m + \eta_{ns}$. Also, the difference between $\bar{T}_m$ and $\bar{T}_{ns}$ is likely to be greater during observations than during calibration.

For example,

$$160.69^\circ = \frac{1}{(0.98)(0.8674)} [245^\circ - 240^\circ(0.98)(0.0816) - 4.2^\circ(0.98)(0.0510) - 200^\circ(0.02)]$$

<u>Error</u>	<u>$\Delta \bar{T}_m(^{\circ}\text{K})$</u>
$(\eta_m + \eta_{ns})$	1.65°
$\eta_m \downarrow, \eta_{ns} \uparrow$	0.45°
k_r	0.24°
T_A	0.82°
$\bar{T}_{ns}$	0.47°
$\bar{T}_o$	0.47°

The total system temperature accuracy is found to be $\pm 2.0^\circ \text{K}$.

References

1. P. A. R. Ade, J. D. G. Rather, and P. E. Clegg, "Limits to Solar Limb Darkening at a Wavelength of 1.4 Millimeters Derived from Antenna-Beam Parameters," *Astrophysical Journal*, Vol. 187, 15 January 1974, pp. 389-392.
2. J. W. M. Baars, "The Measurement of Large Antennas with Cosmic Radio Sources," *IEEE Transactions on Antennas and Propagation*, Vol. AP-21, No. 4, July 1973.
3. R. H. Harten, "Beam Characteristics of the 300-Ft. Telescope," *Astronomical Journal*, Vol. 78, No. 7, September 1973.
4. N. J. Keen, "A Note on Radiometric Noise," *Microwaves*, October 1976.
5. J. D. Kraus, Radio Astronomy, McGraw-Hill, New York, 1966.
6. A. D. Kuzmin, and A. E. Salomonovich, Radioastronomical Methods of Antenna Measurements, Academic Press, New York, 1966.
7. G. S. Levy, D. A. Bathker, Ludwig, D. E. Neff, and B. L. Seidel, "Lunar Range Radiation Patterns of a 210-Foot Antenna at S-Band," *IEEE Transactions on Antennas and Propagation*, 1967.
8. J. Ruze, "Antenna Tolerance Theory - A Review," *Proceedings of the IEEE*, Vol. 54, No. 4, April 1966.
9. B. L. Ulich, J. R. Cogdell, and J. H. Davis, "Planetary Brightness Temperature Measurements at 8.6 mm and 3.1 mm Wavelengths," *Icarus* 19, 1973, pp. 59-82.
10. B. L. Ulich, "Absolute Brightness Temperature Measurements at 2.1-mm Wavelength," *Icarus* 21, 1974, pp. 254-261.

11. B. L. Ulich, and R. W. Haas, "Absolute Calibration of Millimeter-Wavelength Spectral Lines," Astrophysical Journal Supplement Series, Vol. 30, March 1976, pp. 247-258.
12. B. L. Ulich, private communication.
13. M. E. Viggh, "How Noisy Is That Load?," Microwaves, January 1976.

5. Antenna Interface

5.5 SYSTEMATIC ERRORS

System noise temperature variation with frequency due to radiated mixer noise reflected from the antenna feed mismatch places severe constraints on subreflector position stability. These constraints can be eased using optical design procedures to greatly reduce the mismatch between the antenna and the feed.

A millimeter wave mixer radiates noise power at an elevated temperature of $\approx 400^\circ\text{K}$ due to nonthermal noise generation in the diode. Some of this energy is reflected back into the mixer by the small mismatches at the chopper wheel reflecting surface and at the antenna subreflector. These reflections cause variations in system noise temperature with frequency. The period of this variation is

$$\Delta f = \frac{c}{d}$$

where c is the speed of light and d is the two way reflection path length. For the MASR antenna system the period of the antenna subreflector standing wave is 88 MHz.

Shown in Figure 1 is a system temperature calibration of the National Radio Astronomy Observatory 11 meter millimeter wave antenna operating as a single sideband spectrometer at 88.6 GHz. System temperature variations with two distinct periods are evident. The rapid variation with a period of 22.8 MHz is due to the reflection from the antenna subreflector, and the larger variation with a period of 280 MHz is due to the reflection from the chopper wheel absorbing surface. System temperature variation in the MASR due to chopper wheel reflections will be less than shown here due to the use of a higher performance absorber.

The system temperature of each radiometric channel in the MASR is determined during calibration at the end of every frame. These reflections will affect temperature accuracy only if the reflection path length changes between calibrations, shifting the phase of the standing wave in the IF. This effect is most noticeable in the narrow IF channels which only cover part of a variation period. The reflection distance from the feed horns to the chopper surface is not likely to vary, and this reflection will not degrade system accuracy. However, the reflection path length to the main subreflector must be accurately controlled. With a subreflector mismatch similar to that present in the NRAO telescope, the subreflector position must be stable to within ± 1 mil between calibrations to keep the maximum temperature error within bounds. An estimated 0.7°K error will result in some channels with a subreflector position change of 1 mil. The error in the rest of the system will be less.

Computer programs developed to design optical systems can be used to reduce the amount of radiated energy reflected back into the feed and ease the subreflector stability requirement. LACOMA is an optical design program developed at Hughes which emphasizes the reduction of backscatter from optics and structure.

Commercially available programs such as ACCOS-5 and POLYPAGOS are currently used in the design of direct detection optical systems where baffles are required to isolate the detector from scattered light. The use of these powerful design aids will allow the subreflector mismatch to be greatly reduced.

References

1. S. Silver, Microwave Antenna Theory and Design, Radiation Laboratory Series, Volume 12, McGraw Hill, New York, 1949, p. 443.
2. B. L. Ulich, private communication.

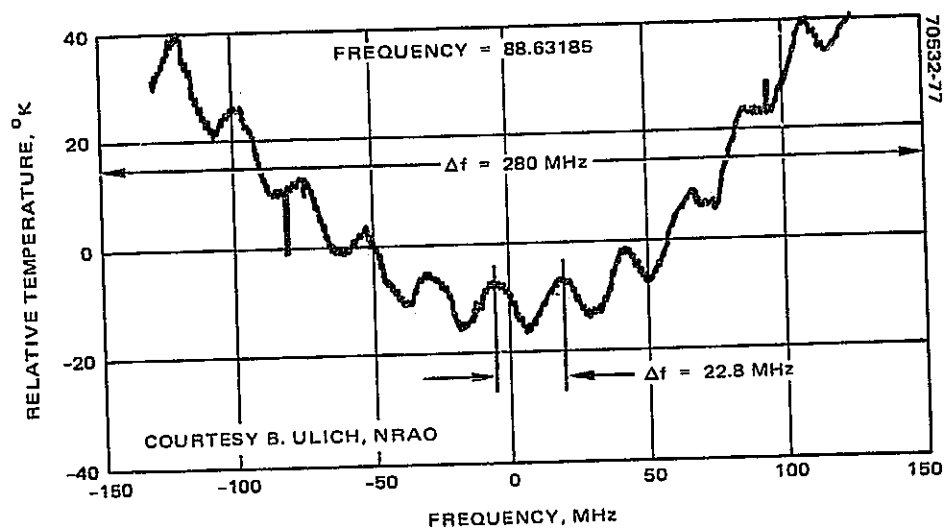


FIGURE 1. RELATIVE SYSTEM TEMPERATURE VERSUS FREQUENCY

APPENDICES

A	MASR CONCEPT	A-1
B	REFERENCE AVERAGING IN PRESENCE OF GAIN FLUCTUATIONS	B-1
C	LASER BEAMS AND RESONATORS	C-1
D	PROPERTIES OF MATERIALS FROM 100 TO 200 GHz	D-1

APPENDIX A. MASR CONCEPT

1) MISSION REQUIREMENTS

The mission of the microwave atmospheric sounding radiometer (MASR) is to collect data to aid in the observation and prediction of severe storms and other mesoscale weather phenomena. The MASR is to be launched via shuttle into a geostationary altitude orbit. This orbit choice allows the continuous measurement needed to resolve mesoscale dynamics. The instrument will operate in conjunction with multispectral imaging and infrared sounding sensors, and its function is to provide three dimensional temperature and humidity fields, primarily in cloudy and overcast atmospheric regions. Further mention of the optical sensors will be limited to noting that the other sensors have stringent line of sight pointing requirements. This will limit allowable spacecraft attitude perturbations due to reaction torques of scanning the MASR. A 20 minute pitch and roll stability of 11 μ rad is typical of the requirements of the advanced atmospheric sounding and imaging radiometer (AASIR).

As part of an IR&D program, Hughes has been assessing the observational requirements for satellite data for all scales of weather phenomena. An understanding of the data requirements for mesoscale weather prediction has been derived from numerous discussions with NASA, NOAA, military, and university personnel, as well as from the literature in the field. The requirements listed below represent Hughes current understanding of how the MASR might be used in a severe thunderstorm environment, but obviously represent only one of a number of scenarios for the use of the MASR. This material is included to indicate Hughes continuing interest in mesoscale forecasting problems.

The observation requirements for mesoscale weather phenomena are as varied as the phenomena themselves. A particularly simple mesoscale phenomenon to predict is frost, in which knowledge of the synoptic conditions, surface temperature, and existing and projected cloud cover is adequate to forecast frost a few hours in advance. At the other extreme are the data requirements for the severe thunderstorm environment, where knowledge of the synoptic environment, coupled with frequent measurements from the synoptic scale down to the cloud droplet scale, are needed. The severe thunderstorm has been chosen to size the MASR mission requirements, both because it places severe time constraints on frame repetition rates and because of the economic and social benefits to be derived from the improved prediction of this phenomenon.

What are the measurement requirements imposed by the severe thunderstorm environment? The time and space scales over which measurements are needed depend upon the state of the thunderstorm development. In the early stages of the storm, before convergence and convection have begun, knowledge of the temperature field with synoptic scale resolution (≈ 400 km in the horizontal and ≈ 8 levels in the vertical) and the humidity field and location of the clouds with mesoscale resolution (≈ 20 to 30 km in the horizontal and ≈ 10 vertical levels), with a measurement frequency of once per 2 hours, are the measurements of primary importance. After convergence has begun and the inversion layer begins to rise, all variables should be measured with mesoscale resolution and with a once per hour frequency. Finally, from the beginning of cumulus cell development until the end of the storm, measurements with mesoscale horizontal and vertical resolutions should be made approximately every 15 minutes. At the same time, measurements with sub-mesoscale resolution and 15 minute repetition rates would be desirable for precise location of storm activity, extent of storm development, and for forecast verification. Submesoscale resolution from satellite data will be limited in the near future to visible and infrared imagery data.

The use of the MASR should be limited to measurements in cloudy and partly cloudy atmospheric regions. Its coverage rate of 750 km^2 in 30 minutes means that it would take nearly 6 hours to make a measurement over the 2500 km^2 extent of a typical numerical prediction model, which might be used in the early morning to identify areas of probable storm activity in the afternoon and evening. Rather, it should be used to sample temperature and humidity distributions in and around selected cloudy regions of the 2500 km^2 grid structure, and sampling should be completed in about 1 hour.

The coverage area of the MASR should be further limited, and its repetition rate increased as thunderstorm development proceeds throughout the day. After the beginning of cumulus development, the MASR will be most useful in measuring temperature and humidity under the cloud top level. During the height of the storm, the MASR will be required to sound beneath the thunderstorm's cirrus anvil cover and may provide semiquantitative precipitation data in the center of the thunderstorm over an area of 500 km^2 with a 15 minute repetition rate. Estimates of the accuracy requirements for the temperature and humidity fields are between 1° and 2° K for the temperature field and approximately 10 percent for the absolute humidity field.

In addition to the quantitative uses of MASR data, it is likely there will be a desire to produce temperature and humidity maps of these data for visual analysis and interpretation. A possible application of this type would be to map total vertically integrated water vapor as a function of time. Areas where the water vapor is increasing would indicate areas of convergence and likely locations for future convective overturning. Another application of this type would be to combine both the temperature and humidity sounding data to produce a stability map using one or more of the conventional stability indices. The purpose again would be to identify areas of convective instability. In conclusion, a summary of the performance needed for the MASR to meet its mission objectives is given in Table 2-1.

TABLE 2-1. MASR MISSION REQUIREMENTS

• Spatial resolution	
Horizontal	20 to 30 km
Vertical	8 to 10 levels up to 100 mb
• Coverage rate	750 km ² /30 min
• Temperature accuracy	1° to 2°K
• Humidity accuracy	10 to 20% absolute

2) SYSTEM REQUIREMENTS

The horizontal spatial resolution requirements for temperature and humidity sounding of severe storms vary for different severe storms and meso-scale phenomena, but most can be satisfied with the 20 to 30 km resolution element indicated in Table 2-1. For nadir observation from synchronous altitude, this range corresponds to antenna beam widths in the range 0.56 to 0.84 mr, or 0.032° to 0.048°.

Due to the fundamental limitation of diffraction, a typical radiometer antenna of diameter D can only give a beamwidth of the order $1.3 \lambda/D$, where λ is a wavelength. The frequencies of the microwave absorption resonances of O_2 and H_2O , which are available for the sounding of atmospheric temperature and water vapor, respectively, are shown in Figure 2-1. The 60 GHz O_2 and 22 GHz H_2O resonances have been used in microwave sounders flown or designed to date, since these are for observing synoptic scale phenomena from low altitudes (~1000 km), where antenna beamwidths in the range 7° to 10° are adequate. For geosynchronous sounding of severe storms, there is a need to use the higher frequency O_2 resonance at 118 GHz (2.53 mm) for temperature sounding, and the H_2O resonance at 183 GHz (1.64 mm) for water vapor sounding. These shorter wavelengths allow the required narrow beamwidths to be achievable with practical antenna sizes (<4.4 m) that can fit into the shuttle payload bay without furling.

The need for a narrow antenna beam thus leads to two critical requirements for geosynchronous microwave sounding:

- Large aperture reflector antenna, 3 to 4.4 m in diameter
- Low noise radiometer operating between 118 and 183 GHz

To map a finite area on the earth's surface, the antenna beam has to scan in two orthogonal directions. Because of the very high frequencies utilized and the requirement for low loss reception, electrical scanning is impractical to implement and mechanical scanning is required. This leads to a third critical requirement:

- Mechanical scanning of a large antenna in two orthogonal axes.

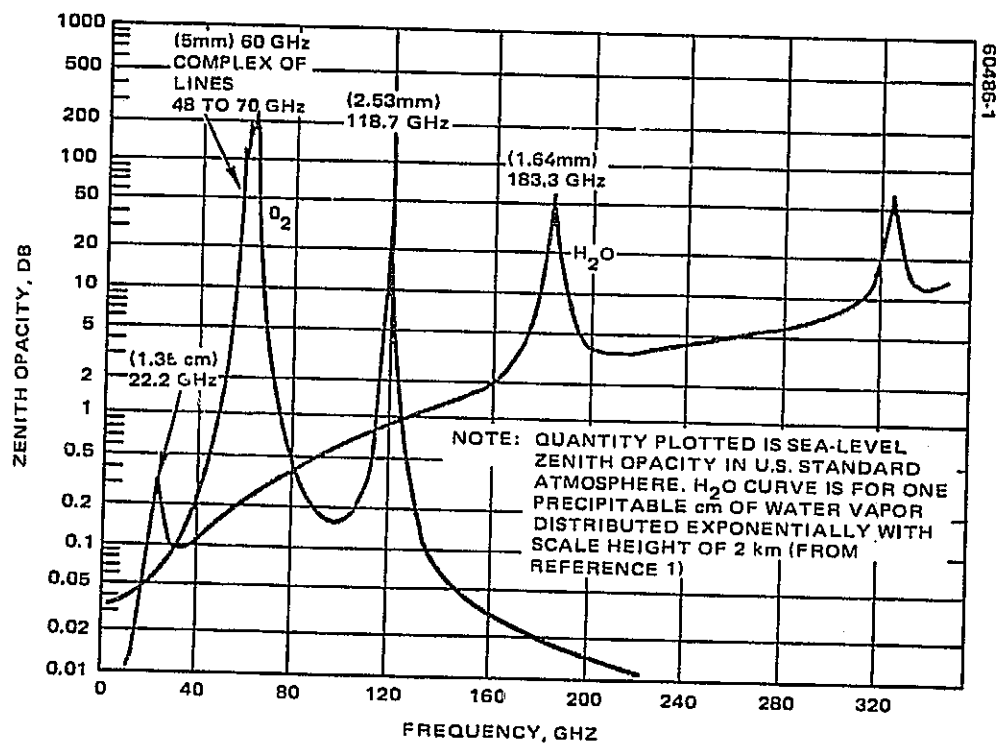


FIGURE 2-1. MICROWAVE ABSORPTION RESONANCES OF O_2 AND H_2O AVAILABLE FOR SOUNDING

This last requirement is subject to the constraints that the antenna scan motion must not cause excessive reaction torques on the spacecraft that would violate the pointing accuracy and stability requirements of other instruments (e. g., AASIR).

These three requirements determine the general characteristics of a microwave sounding system needed for observing severe storms from geosynchronous altitude. The objective of the two proposed studies is to verify the feasibility of developing such a system for a future flight in the 1982-1983 time frame.

3) MASR REQUIREMENTS

IF Bandwidth and Channel Selection

In order to better appreciate the basis for the system requirements involved in the selection of operating frequencies and IF bandwidths, it is necessary to review briefly the principle of microwave sounding and radiative transfer.

The MASR is a multichannel radiometer (or spectrometer) for performing microwave sounding experiments from a geosynchronous satellite. Microwave sounding consists of measuring, at selected frequencies, the brightness temperature of passive microwave radiation emitted by the earth's atmosphere and surface, and the derivation of meteorological parameters from these measurements. The basic equation that relates the brightness temperature of the earth's outgoing radiation to the atmospheric parameters is the microwave radiative transfer equation.

$$T_B(f) = \epsilon_s T_s e^{-\int_0^\infty \alpha(f, z) dz} + \int_0^\infty \alpha(f, z) e^{-\int_z^\infty \alpha(f, z') dz'} T(z) dz \quad (1)$$

$$+ (1 - \epsilon_s) e^{-\int_0^\infty \alpha(f, z) dz} \cdot \int_0^\infty \alpha(f, z) e^{-\int_0^z \alpha(f, z') dz'} T(z) dz$$

where

$T_B(f)$ = brightness temperature of earth's outgoing microwave radiation as a function of frequency f

$T(z)$ = atmospheric temperature as a function of altitude z

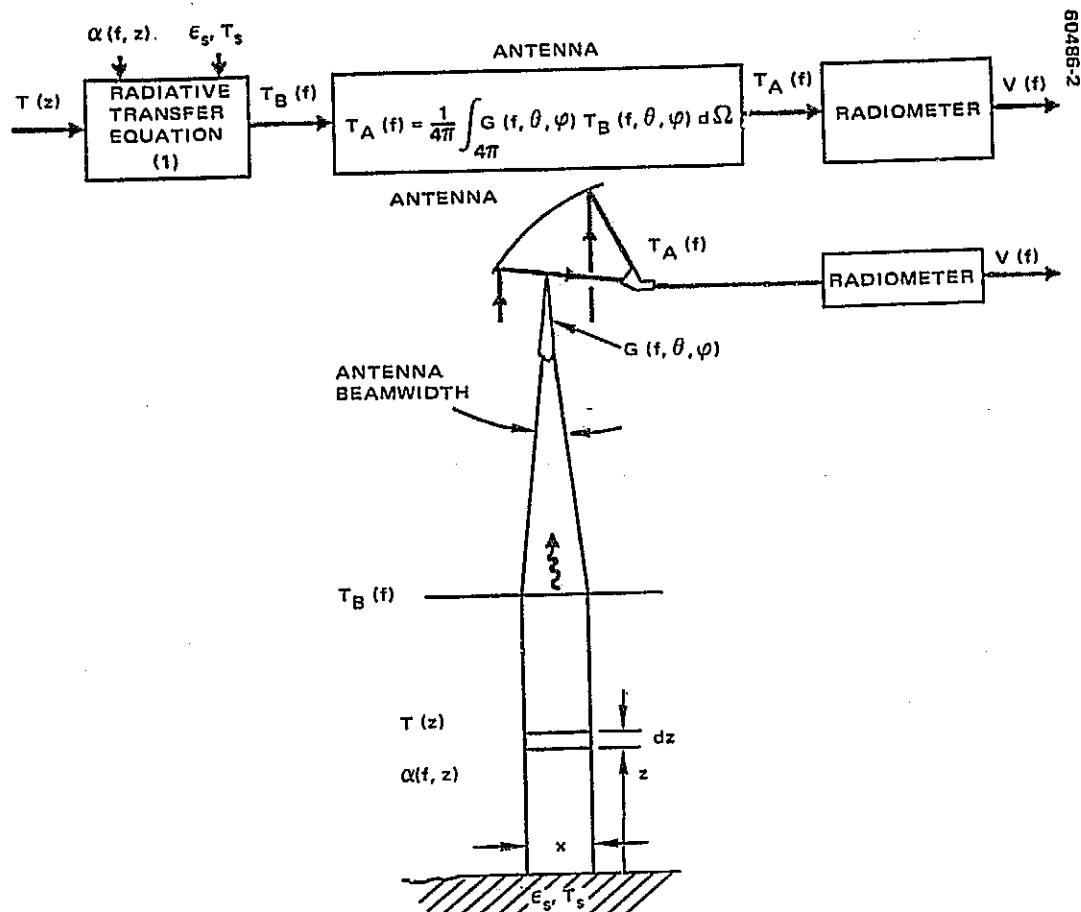


FIGURE 2-2. MICROWAVE SOUNDING SYSTEM FUNCTIONAL FLOW DIAGRAM

ϵ_s and T_s = emissivity and temperature, respectively, of earth's surface
 $\alpha(f, z)$ = absorption coefficient of atmosphere for microwaves of frequency f ; varies with altitude

$$\alpha(f, z) = \sum_i k_i [f, p(z), T(z)] n_i(z) \quad (2)$$

where the sum is over the number of molecular species that contribute to absorption at frequency f , $n_i(z)$ is the i th absorber number density profile, and $k_i(f, p, T)$ is the molecular absorption coefficient of the i th absorber, which usually depends strongly on pressure p but weakly on temperature T . Equation 1 is for vertical (nadir) viewing as illustrated in Figure 2-2. For off nadir viewing, a similar equation can be written for an inclined atmospheric column.

The physical significance of the different terms in Equation 1 are as follows: The first term is the surface emission, attenuated by the atmosphere. The second term is a summation of emissions from all the layers of the atmospheric column (Figure 2-2), the contribution from each layer being attenuated by the intervening atmosphere between the layer and outer space. The third term is due to surface reflection effect. It accounts for downward atmospheric radiation reflected by the surface, and which eventually contributes to the outgoing radiation.

The theory for the different microwave sounding experiments is generally as follows: In temperature sounding, ϵ_s , T_s , and $\alpha(f, z)$ are assumed known, then the temperature profile $T(z)$ is inferred from measurements of $T_B(f)$, the two functions being related by Equation 1. In humidity sounding, ϵ_s , T_s and $T(z)$ are assumed known, and $\alpha(f, z)$ is adjusted in Equation 1 by varying the water vapor concentration until the observed and calculated values of $T_B(f)$ agree. In window measurements, the surface term in Equation 1 dominates over the other two terms, and the measured brightness temperature can be used to infer surface temperature or emissivity.

Figure 2-2 shows a functional flow diagram for a microwave sounding system. The brightness temperatures at the selected frequencies are the primary quantities related to the atmospheric parameters. However, the radiometer measures directly the antenna temperature at the selected frequencies. The antenna temperature is related to the brightness temperature from all directions by the equation

$$T_A(f) = \frac{1}{4\pi} \int_{4\pi} G(f, \theta, \phi) T_B(f, \theta, \phi) d\Omega$$

where $G(f, \theta, \phi)$ is the antenna gain function, (θ, ϕ) are polar and azimuthal angles about the antenna. The above equation together with the radiative transfer Equation 1 form the system equations shown in Figure 2-2.

Rewriting Equation 1 in a more compact form

$$T_B(f) = \epsilon_s T_s e^{-\tau(f)} + \int_0^{\infty} W(f, z) T(z) dz \quad (3)$$

where

$$\tau(f) = \int_0^{\infty} \alpha(f, z) dz$$

is the zenith opacity and $W(f, z)$, which depends on $\alpha(f, z)$, is a weighting function that weighs the contribution to T_B from the atmospheric temperature at various altitudes. For the case where the surface reflectance is negligible

$$W(f, z) = \alpha(f, z) e^{-\int_z^{\infty} \alpha(f, z) dz} \quad (4)$$

The weighting function usually peaks at some altitude and this altitude varies with frequency. In temperature sounding, $T_B(f)$ is measured at several discrete frequencies f_i about an O_2 absorption line, each value of f_i being selected for the desired altitude where the weighting function $W(f_i, z)$ will peak.

A consideration in choosing the set of $W(f_i, z)$ is that their mutual overlap be minimized, since overlapping leads to interdependence among the different $T_B(f_i)$ measurements and is responsible for the instability in inverting Equation 3 (Reference 2). To minimize overlap, the half-width of each $W(f_i, z)$ should be minimized, and the frequencies f_i be chosen to space the $W(f_i, z)$ over the altitude range to be covered.

At each RF, what is measured is the average brightness temperature $\overline{T}_B$ over a spectral band $\Delta f = B_{IF}$. The appropriate weighting function for $\overline{T}_B$ is a band averaged weighting function

$$\overline{W}(f, z) = \frac{1}{\Delta f} \int_{\Delta f} W(f, z) df \quad (5)$$

In general, as Δf increases, $\bar{W}(f, z)$ becomes broader in half-width. Therefore, it is desirable to keep Δf or B_{IF} small enough so that $\bar{W}(f, z)$ will approach the monochromatic $W(f, z)$, which has the minimum width. Hence, the maximum value of B_{IF} for the temperature sounding channels is limited by the need to maintain vertical resolution for the sounder.

To each IF in Table 1 of the radiometer RFP, there corresponds two RF symmetrical about the 118.75 GHz O₂ line center. To each RF frequency, there corresponds a weighting function $W(f, z)$ that can be calculated from the known oxygen absorption characteristics about 118.75 GHz and the known oxygen concentration profile in the atmosphere. However, if $\alpha(f, z)$ is assumed to be symmetrical about 118.75 GHz, then $\tau(f)$ and $W(f, z)$ are symmetrical, which leads to $T_B(f)$ being symmetrical as well. Each MASR temperature sounding channel will observe the average of the brightness temperatures at $118.75 - f_{IF}$ and $118.75 + f_{IF}$, and there is effectively only one weighting function for each of the T-channels. The values H_{ave} shown in Table 1 of the radiometer RFP indicate the average height of each channel weighting function. Channels T-1 to T-4 are designed to probe the troposphere, while the other channels are to probe the lower stratosphere. Figure 2-3 shows some calculated weighting functions on the low frequency side of the 118.75 GHz line, which are taken from Reference 3. A few of these weighting functions are at frequencies approximately the same as some of the lower sideband frequencies in Table 1 of the RFP, and the average heights of the weighting functions in Figure 2-3 generally agree with the corresponding average heights given in Table 2-1.

Factors Affecting System Performance

The Figure 2-4 diagram summarizes how the various system requirements interact and lead to the radiometer performance needed. The requirements of area coverage, temporal and (horizontal) spatial resolution of severe storm phenomena, combine together to determine the dwell time available for sounding each FOV cell. Contiguous coverage of an area L km by L km within a frame time t_F , with a nadir ground resolution size of X km, leads to an FOV dwell time, or integration time

$$t_i \approx \frac{X^2}{L^2} t_F$$

This formula is approximate in that it neglects the time to turn around, step, and calibrate at the end of each scan, and it assumes that the $L \times L$ area is located in the subsatellite region. Contiguous coverage of the 183 GHz channels is required to generate water vapor maps. This requirement leads to an integration time of $(t_i)_{183} \sim 1$ second as specified in Table 2 of the RFP. This value is obtained from the above equation with $X = (35,800 \text{ km}) (1.3 \lambda/D)$, if we substitute in the maximum antenna diameter of 4.4 m, and $L = 750 \text{ km}$, $t_F = 30 \text{ min}$ as typical mission requirements (Table 2-1). The

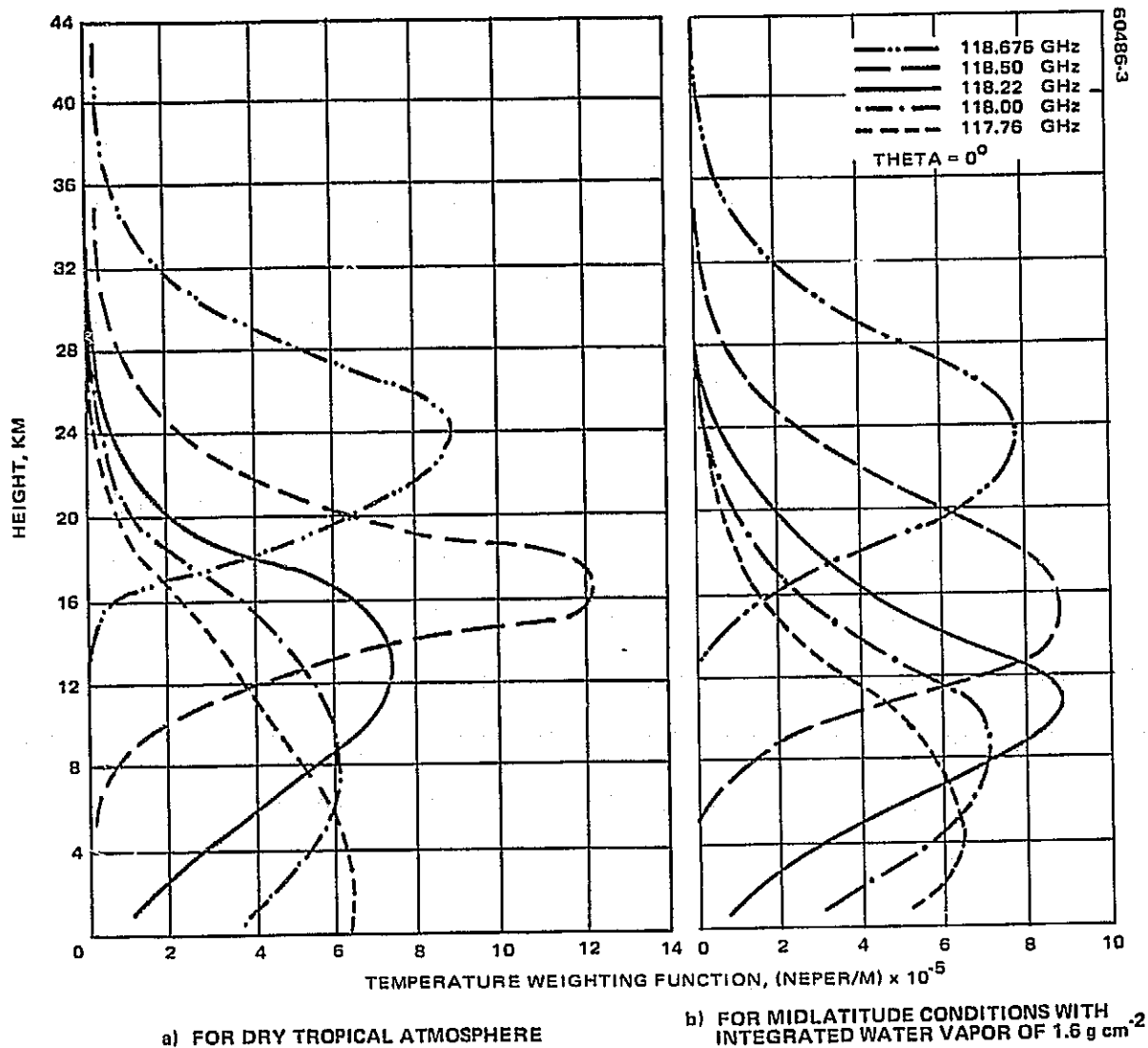


FIGURE 2-3. TEMPERATURE WEIGHTING FUNCTIONS FOR CHANNELS NEAR 118 GHz LINE (FROM REFERENCE 2)

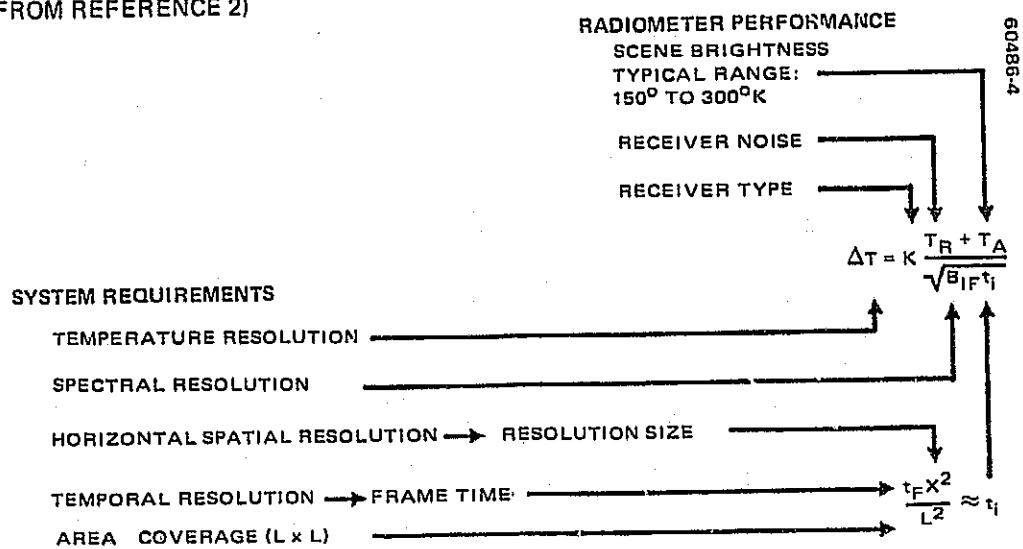


FIGURE 2-4. SUMMARY DIAGRAM RELATING SYSTEM REQUIREMENTS TO RADIOMETER PERFORMANCE

integration time and beamwidth for the 183 GHz channel determine a scan rate $\dot{\theta}_{sc}$, at which rate the antenna scan mechanism must be designed to move the 183 GHz beam. But the same mechanism will move the other channel beams at the same rate, so the FOV dwell times for the other channels are simply determined by $(t_i)_\lambda = (1.3 \lambda/D)/\dot{\theta}_{sc}$, or $(t_i)_\lambda$ is proportional to λ . Hence, $(t_i)_{118}$ is $183/118 = 1.55$ times $(t_i)_{183}$, as indicated in Tables 1 and 2 of the RFP.

As indicated in Figure 2-4, the requirements of temperature and spectral resolution determine the required ΔT and B_{IF} , respectively. These requirements differ for different types of sounding experiments. For temperature profile sounding, the accuracy required in measuring $T_B(f)$ has to be typically many times the accuracy being sought for the temperature profile. This is because the mathematical "inversion" techniques used to invert the integral Equation 1 to recover $T(z)$ are highly sensitive to data noise, as the computations deal with inversion of ill-conditioned matrices. For example, in Tiros-N, where temperature profiles accurate to 1° to $1.5^\circ K$ are sought, the microwave sounding unit operating at the 60 GHz O_2 band has a $\Delta T = 0.3^\circ K$. Similarly, NEMS at 60 GHz had a $\Delta T = 0.2^\circ$ to $0.3^\circ K$, and was designed to retrieve atmospheric temperatures to accuracies of 1° to $3^\circ K$. The random errors in $T_B(f)$ that can be tolerated is also dependent on the particular inversion algorithm to be used. The ΔT_B accuracy needed can be found through "computational experiments" by simulating the inversion process with artificial noise injected into the data. From the ΔT_B accuracy, one can specify the noise equivalent ΔT_A desired.

The selection of B_{IF} for each temperature sounding channel is a compromise between the desire of increasing B_{IF} to reduce ΔT versus that of decreasing B_{IF} to provide sharper weighting functions. In temperature sounding, spectral resolution is related to vertical spatial resolution; so the B_{IF} tradeoff is between temperature and vertical resolution.

In water vapor sounding, the required ΔT is controlled by the ~10 to 20 percent accuracy desired in measuring the water vapor profile $n_{wv}(z)$. The question of what random errors in the set of 183 GHz $T_B(f_i)$'s will correspond to 10 to 20 percent error in $n_{wv}(z_i)$'s is dependent on the numerical algorithm used to go from the set of $T_B(f_i)$'s to the set of $n_{wv}(z_i)$'s. A way to answer this is again to perform computational experiments on the actual algorithm to be used, by injecting artificial noise into the water T_B data vector.

APPENDIX B. REFERENCE AVERAGING IN PRESENCE OF GAIN FLUCTUATIONS

1. Introduction

The temperature resolution of a microwave radiometer is determined by errors from two sources: (1) random fluctuations inherent in the square-law detection of noise power in a frequency band; and (2) temporal variations in gain. The use of "reference averaging" to improve a Dicke radiometer, as first proposed by Bremer and Prince,⁽¹⁾ has the effect of reducing the first type of errors. The technique reduces the variance in the reference voltage by increasing its integration time for postdetection smoothing, which leads to a decrease in the overall variance of the differential output voltage. In a switching radiometer with 50% duty cycle, instead of integrating over one-half the dwell time (τ_p) for a resolution element or "pixel", we can integrate over $N\tau_p/2$ by averaging the reference samples in N consecutive pixels. Such an extension of integration time is possible for the reference samples because the reference temperature is essentially constant. On the other hand, the integration time for the antenna samples must be limited to one-half the pixel dwell time in order to resolve the variation of scene temperature from pixel to pixel.

As the reference averaging time $N\tau_p/2$ increases, the temperature error due to detector fluctuations decreases, which corresponds to a decrease in the K-factor. If the chopping fraction remains at 50%, it can be shown that

$$K = \sqrt{2 + \frac{2}{N}} \quad (1)$$
 Mathematically, as $N \rightarrow \infty$, $K \rightarrow \sqrt{2}$, which corresponds to about 30% reduction in K from its original value of 2. However, physically, it is not clear how large a value of N can or should be used, which would determine how much improvement is physically realizable. Before this question can be answered, it is necessary first to understand how reference averaging affects the second type of errors; in particular, to know also how the errors due to gain variations depend on the parameter N . This paper is aimed at clarifying these questions. Our results will show that the gain variation error increases with increasing N , so there is, in fact, an optimal value of N that minimizes the combination of detector fluctuation and gain variation errors.

Above, we have described reference averaging with symmetric chopping. A further reduction in detector noise is possible by incorporating asymmetric chopping with reference averaging. If the fraction of time for antenna sampling is increased from .5 to some value q ($.5 < q < 1.0$), the antenna integration time is lengthened from $.5\tau_p$ to $q\tau_p$, while the reference integration time in each pixel is shortened from $.5\tau_p$ to $(1-q)\tau_p$. However, the latter effect is counteracted by averaging the references over N pixels to give a total reference integration time of $N(1-q)\tau_p$. With a large enough N , the overall effect is to reduce both the variances of the reference and antenna voltages. The K -factor now becomes a function of N and q , and

is expressed by

$$K = \sqrt{\frac{1}{q} + \frac{1}{N(1-q)}} \quad (1)$$
 For a given N , K is minimum when

$$q = q_{op} = \frac{\sqrt{N}}{\sqrt{N} + 1}$$
 . If q is set equal to q_{op} for each N value, then as

N is increased, the K-factor decreases as $K = 1 + \frac{1}{\sqrt{N}}$ (1). Thus, as $N \rightarrow \infty$,

$K \rightarrow 1$, which is the value for a total power radiometer. As in the symmetric chopping case, one might again inquire how large a value of N can be used, in an attempt to ascertain the maximum improvement achievable. However, in any practical situation, there will always be gain variabilities, in which case the question of maximum improvement must be reformulated differently. We already know the detector fluctuation error as a function of N and q, but we also need to know the gain variation error as a function of N and q, so we can determine the values of N and q that will minimize the combined error due to both effects. This will then give the maximum resolution improvement achievable by reference averaging with asymmetric chopping.

In this paper, we will start with analyzing only the case of reference averaging with symmetric chopping. Our approach is to start with a conventional Dicke system with symmetric square-wave chopping and square-wave detection, and examine what happens to the gain variation error as reference averaging is incorporated with increasing values of N. The chopping fraction is kept constant at 50%, while N is varied. This facilitates comparison with the conventional Dicke case, and allows easy interpretation of the results. The effect of adding asymmetric chopping to reference averaging will be considered in later studies. However, some of the characteristics we will derive for the symmetric chopping case will be qualitatively similar to that for asymmetric chopping, though the details will differ; for example, the combined error due to detector fluctuation and gain variation when plotted against N will show a minimum, for any constant value of q, similar to what we derive for $q = .5$.

The effect of gain variabilities on a reference averaging radiometer was discussed in Reference 1, but was not treated with adequate depth to

reveal the tradeoff between gain variation and detector noise errors. Consequently, there was some lack of appreciation of the limitation on the improvement realizable from reference averaging.

Ross Graves⁽²⁾ has studied the problem of reference averaging in the presence of gain variation as a statistical estimation problem; i.e., as a problem in obtaining a best estimate to the true value of the time-varying reference voltage at a specified time (the observed pixel time), given data samples of the reference voltage at other times which are contaminated by the detector noise. He treated the generalized problem of using N pixels of reference samples to estimate an m -parameter reference-voltage function of time ($m < N$), and studied also the generalized technique of reference averaging with non-uniform weighting of the samples. His study utilized very elegant techniques to arrive at a number of interesting results, and the reader is referred to Reference 2 for a discussion of his numerous results.

In this paper, we will examine the reference averaging system in its originally proposed form, where the one-parameter estimator is a constant, equal to the unweighted, simple average of the reference samples over N pixels. "Symmetric" averaging is assumed, where the N consecutive pixels of reference samples to be averaged are time-symmetrically distributed about the center of the observed pixel dwell time. Noting that gain variations are independent of detector fluctuations, we may set the latter to zero while we analyze the errors due solely to gain variations. This is the reverse of setting the gain variations to zero while analyzing the errors due solely to detector fluctuations, as was done in deriving the K -factor improvement for ideal reference averaging. The errors derived for these two opposite cases may be combined by RSS to yield the radiometer's temperature resolution. Since the detector fluctuation noise is "turned off", we may leave aside the statistical

estimation aspects of the problem for the time being, and concentrate on studying the signal processing operations implied by the reference averaging algorithm. We want to find out how the radiometer output response to the antenna temperature signal is modified by gain variation with time which may be deterministic and/or random; and how these modifications depend on the reference averaging system parameters, such as N , τ_p and the chopping period. The objective is to relate antenna temperature error to gain variabilities, and to examine the dependence of that relationship on the parameters of reference averaging. Another objective is to derive equations that can be used for future computer simulation of reference averaging systems in general, to study system parameter tradeoffs and error sensitivities.

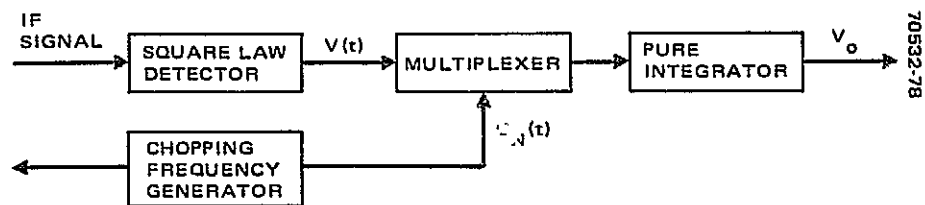


FIGURE 1. POSTDETECTION BLOCK DIAGRAM OF REFERENCE AVERAGING RADIOMETER

2. Response of Dicke vs. Reference Averaging Radiometer

To compare the reference averaging and Dicke systems, it is helpful to cast both signal processing schemes into a common format, so that they can be readily compared. Toward this end, we view both systems as having postdetection sections equivalent to the block diagram of Figure 1. The modulated output from the square-law detector, denoted by $V(t)$, is multiplied by a reference waveform $C(t)$ that is synchronized with the modulation frequency and phase. The product $C(t)V(t)$ is filtered and smoothed by a pure integrator, to yield the radiometer output V_o . The multiplier reference waveform $C(t)$ is sometimes referred to as the "correlation" in Dicke radiometry.⁽³⁾ It will be shown that by properly defining the correlation $C(t)$ a digital reference averaging system can be made equivalent to the system of Figure 1. The function $C(t)$ is different for different values of N , and we will study the sequence of $C_N(t)$'s.

The modulated output from the square-law detector is given by

$$V(t) = Bkg(t)[T(t) + T_{RN}] + n(t) \quad (1)$$

where B is the IF or predetection bandwidth, k is Boltzmann's constant, $g(t)$ is the power gain, T_{RN} is the receiver noise temperature, $T(t)$ is the noise temperature at the switching terminal, and $n(t)$ is the zero-mean random process representing the fluctuations in the square-law detector. Later on, $n(t)$ will be set equal to zero, when we analyze the errors due solely to gain variations. The notation is simplified by defining the product

$$G(t) = Bkg(t) \quad (2)$$

and rewriting $V(t)$ as

$$V(t) = G(t)[T(t) + T_{RN}] + n(t) \quad (3)$$

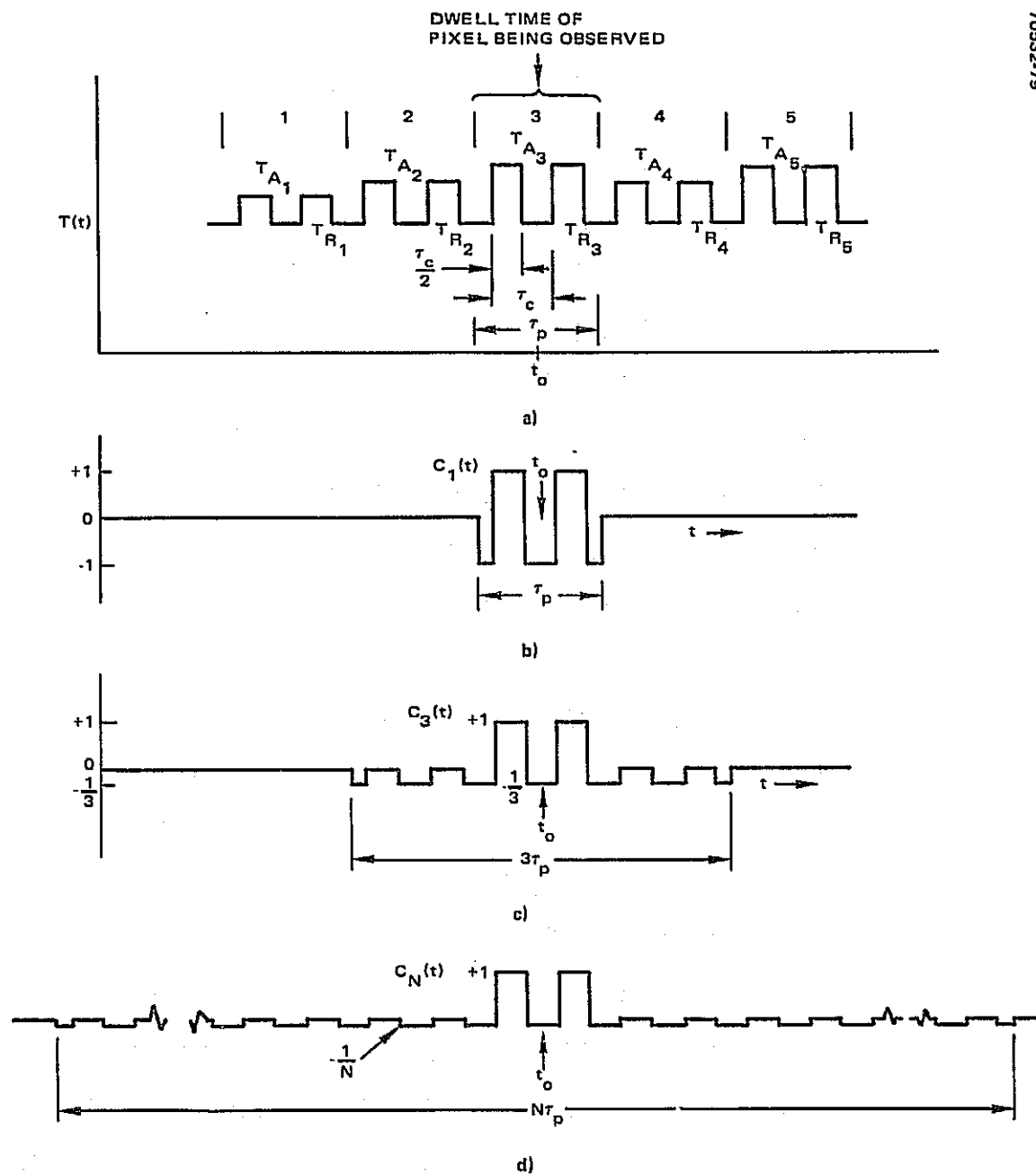


FIGURE 2. RADIOMETER OUTPUT WAVEFORMS

We will refer to $G(t)$ as the "gain", with the understanding that it is really the product of the electronic gain and kB ; so variations in $G(t)$ can arise from variations in the electronic gain and/or bandwidth. Therefore, the error formulas to be derived for variations in $G(t)$ can be applied to either of these two effects, or a combination of both.

Consider a Dicke radiometer with 50% chopping fraction and a chopping period τ_c . $T(t)$ will be a square-wave function of time that alternately equals the antenna temperature T_A and reference temperature T_R every $\tau_c/2$ seconds. In general, T_A is a function of time as the radiometer beam scans the scene, while T_R is essentially constant, though it may have some residual variation depending on how well the reference source temperature is stabilized. To simplify the analysis, the radiometer beam will be assumed to move from pixel to pixel in an idealized step-dwell mode, with the stepping time being negligible compared to the dwell time. Under this assumption, a typical plot of $T(t)$ might look like that shown on Figure 2(a). The labeled intervals along the top denote pixel dwell time slots, each having a width equal to the pixel dwell time τ_p . For simplicity of illustration, the waveform is drawn for the case $\tau_c = \tau_p/2$. T_A is constant within each pixel dwell time, but its value will generally change from pixel to pixel following the scene brightness variation. The values of T_A and T_R within the j -th pixel time are denoted by T_{A_j} and T_{R_j} , respectively. The corresponding values of $V(t)$ when $T(t)$ equals T_{A_j} and T_{R_j} will be denoted by V_{A_j} and V_{R_j} , respectively.

Referring to Figure (1), the radiometer output V_o has the general form

$$V_o = \text{constant} \times \int C(t) V(t) dt \quad (4)$$

In what follows, the correlation $C(t)$ for the Dicke and reference averaging systems will be determined and compared. In the Dicke case, the radiometer output that indicates the brightness temperature of the i -th pixel (see Figure 2) is simply given by

$$V_o = \bar{V}_{A_i} - \bar{V}_{R_i} \quad (5)$$

where $\bar{V}_{A_i}$ and $\bar{V}_{R_i}$ are, respectively, the average of the antenna and reference segments of $V(t)$, both taken within the i -th pixel dwell time. Combining the integrals in $\bar{V}_{A_i}$ and $\bar{V}_{R_i}$, we can also express V_o as

$$V_o = \frac{2}{\tau_p} \int_{\tau_p} C_1(t) V(t) dt \quad (6)$$

where $C_1(t)$ is a square-wave function that alternates between +1 and -1 every $\tau_c/2$ seconds, as shown in Figure 2(b), and is equal to zero outside the pixel dwell time.

We will use the symbol $C_N(t)$ to denote the correlation for reference averaging over N pixels. Since the Dicke is a special case of the reference averaging system with $N = 1$, the Dicke correlation is denoted by $C_1(t)$. We will find that the $C_N(t)$ functions are easier to describe by sketching them, rather than writing out their algebraic equations; therefore, we will sketch out each function in defining it.

In discussing reference averaging, we will always assume "symmetric" averaging, where the N consecutive pixels of reference samples to be averaged are time-symmetrically distributed about the center of the observed pixel time slot. It is expected that a digital reference averaging system

will most likely use symmetric averaging, since this can be implemented easily with a small amount of memory and can yield great advantage in reducing the system sensitivity to gain variation. For symmetric averaging, if N is odd, we will take an integral number $(\frac{N-1}{2})$ of reference pixels before and after the observed pixel time slot. But if N is even, we will take an integral number plus half of a reference pixel on either side of the observed pixel time.

Consider first the specific case of reference averaging with $N = 3$. The radiometer output that indicates the brightness temperature of the i -th pixel is computed by

$$V_o = \bar{V}_{A_i} - \frac{1}{3} \left(\bar{V}_{R_{i-1}} + \bar{V}_{R_i} + \bar{V}_{R_{i+1}} \right) \quad (7)$$

where $\bar{V}_{A_i}$ is the same as previously defined and $\bar{V}_{R_j}$ is the average of the reference voltages in the j -th pixel. Combining the integrals from all four terms in (7), it can also be expressed as

$$V_o = \frac{2}{\tau_p} \int_{3\tau_p} C_3(t) V(t) dt \quad (8)$$

where $C_3(t)$ is the square-wave function shown in Figure 2(c). It can be described as a train of rectangular pulses of width $\tau_c/2$ and heights equal to $+1$, $-\frac{1}{3}$, or zero, distributed in a symmetric pattern over a time interval of $3\tau_p$. Outside this interval, $C_3(t) = 0$, so its overall width is $3\tau_p$.

Now, consider the case of reference averaging with arbitrary value of N . Assuming that N is odd, the radiometer output for the i -th pixel is given by

$$V_o = \bar{V}_{A_i} - \frac{1}{N} \sum_{j=-m}^m \bar{V}_{R_{i+j}} \quad (9)$$

where $m = \frac{N-1}{2}$. If we again combine the integrals from all the terms, we can write

$$V_o = \frac{2}{\tau_p} \int_{N\tau_p} C_N(t) V(t) dt \quad (10)$$

where $C_N(t)$ for $N > 3$ is illustrated in Figure 2(d). Comparing $C_3(t)$ and $C_N(t)$ for $N > 3$, we note the following: The positive pulses in both functions are located inside the central pixel time slot, and their pulse heights remain equal to unity. The height of the negative pulses in $C_N(t)$ are reduced to a value $-\frac{1}{N}$, while the number of negative pulses is increased, being spread out over a larger interval. The overall width of $C_N(t)$ is increased to a value $N\tau_p$. $C_N(t)$ may be viewed as a train of rectangular pulses of width $\frac{\tau_c}{2}$ and heights equal to $+1$, $-\frac{1}{N}$, or zero, symmetrically distributed over a time interval of $N\tau_p$; outside this interval, $C_N(t) = 0$.

$C_N(t)$ depends not only on N , but also on the chopping period and pixel dwell time. We may indicate this by writing $C_N(t) = C_N(t; \tau_c, \tau_p)$. Substituting Equation (3) into (10), we obtain

$$V_o = \frac{2}{\tau_p} \int_{N\tau_p} C_N(t; \tau_c, \tau_p) G(t) [T(t) + T_{RN}] dt + \frac{2}{\tau_p} \int_{N\tau_p} C_N(t; \tau_c, \tau_p) n(t) dt \quad (11)$$

This generalized system equation can be used to set up a future computer simulation of reference averaging radiometers, to study parameter trade-offs and error sensitivities. It describes the linear response of the

radiometer output to the modulated temperature signal, shows how this response is disturbed by gain variations and detector fluctuations, and indicates how the response is modified by changes in the design parameters N , τ_c and τ_p .

3. Antenna Temperature Errors due to Gain Variability

In what follows, we will set $n(t)$ to zero, and analyze the system behavior in the presence of only gain variations. With $n(t) = 0$, Equation (11) reads

$$V_o = \frac{2}{\tau_p} \int_{N\tau_p} C_N(t) G(t) [T(t) + T_{RN}] dt \quad (12)$$

Consider the above equation for the Dicke case, for which $N = 1$. Within the pixel dwell time, T_A and T_R are constant; so $T(t)$ inside the interval of integration may be represented by

$$T(t) = \frac{1}{2} (T_A - T_R) C_1(t) + \frac{1}{2} (T_A + T_R) \quad (13)$$

where $C_1(t)$ is the Dicke correlation shown in Figure 2(b). Substituting (13) into (12) with $N = 1$, we have

$$V_o = \frac{2}{\tau_p} \int_{\tau_p} G(t) \left\{ \frac{1}{2} (T_A - T_R) C_1^2(t) + \left[\frac{1}{2} (T_A + T_R) + T_{RN} \right] C_1(t) \right\} dt \quad (14)$$

Defining $\bar{G}$ as the average gain within the pixel interval,

$$\bar{G} = \frac{1}{\tau_p} \int_{\tau_p} G(t) dt \quad (15)$$

and γ_1 as the integral

$$\gamma_1 = \frac{2}{\tau_p} \int_{\tau_p} C_1(t) G(t) dt \quad (16)$$

and making use of the fact that $C_1^2(t) = 1$ within the pixel interval, we may rewrite (14) as

$$V_o = \bar{G}(T_A - T_R) + \gamma_1 \left(\frac{T_A + T_R}{2} + T_{RN} \right) \quad (17)$$

The second term in (17) is usually small in a conventional Dicke, but will not necessarily remain small with reference averaging. Therefore, we will retain this term throughout our analysis and will evaluate and compare it for the Dicke and reference averaging case.

If the second term were neglected, as is usually done, we would arrive at the following classic equation for the gain variation error in a Dicke radiometer:

$$\delta T_A = \frac{\delta \bar{G}}{\bar{G}} (T_A - T_R) \quad (18)$$

Evidently, the gain referred to in this familiar formula should be the average gain $\bar{G}$ within the observed pixel dwell time, and $\delta \bar{G}$ is the deviation of the actual $\bar{G}$ at the time of observation from its value at the last calibration.

With the second term present, any variations in $G(t)$ will, in general, cause changes in $\bar{G}$ and γ_1 via Equations (15) and (16). The complete equation for the antenna temperature error caused by gain variations should, therefore, be as follows:

$$\delta T_A = \frac{1}{\bar{G} + \frac{\gamma_1}{2}} \cdot \left[\delta \bar{G} (T_A - T_R) + \delta \gamma_1 \left(\frac{T_A + T_R}{2} + T_{RN} \right) \right] \quad (19)$$

where $\delta \bar{G}$ has the same meaning as above, and $\delta \gamma_1$ is the variation in γ_1 .

Now, consider Equation (12) for the case of reference averaging over N pixels. It will be convenient to divide the integration in Equation (12) into two parts: integration inside the observed pixel where $|t-t_o| < \frac{\tau_p}{2}$, plus integration outside the observed pixel where $|t-t_o| > \frac{\tau_p}{2}$. t_o is the time at the center of the observed pixel interval.

In the region $|t-t_o| < \frac{\tau_p}{2}$, $T(t)$ can be represented by Equation (13) which we rewrite as

$$T(t) = (T_A - T_R) \left[\frac{1}{2} C_1(t) + \frac{1}{2} \right] + T_R \quad (20)$$

where T_A is the antenna temperature in the observed pixel, and $C_1(t)$ is the Dicke correlation. In the region $|t-t_o| < \frac{\tau_p}{2}$

$$C_N(t) \left[\frac{1}{2} C_1(t) + \frac{1}{2} \right] = \frac{1}{2} C_1(t) + \frac{1}{2} \quad (21)$$

Because of this property, the central part of the integral in (12) is given by

$$\begin{aligned} & \frac{2}{\tau_p} \int_{|t-t_o| < \frac{\tau_p}{2}} C_N(t) G(t) \left[T(t) + T_{RN} \right] dt \\ &= (T_A - T_R) \cdot \frac{2}{\tau_p} \int_{|t-t_o| < \frac{\tau_p}{2}} \left[\frac{1}{2} C_1(t) + \frac{1}{2} \right] G(t) dt \\ & \quad + (T_R + T_{RN}) \frac{2}{\tau_p} \int_{|t-t_o| < \frac{\tau_p}{2}} C_N(t) G(t) dt \\ &= (T_A - T_R) \left(\bar{G} + \frac{\gamma_1}{2} \right) + (T_R + T_{RN}) \frac{2}{\tau_p} \int_{|t-t_o| < \frac{\tau_p}{2}} C_N(t) G(t) dt \end{aligned} \quad (22)$$

where $\bar{G}$ and γ_1 are previously defined in (15) and (16).

In the region $|t-t_o| > \frac{\tau_p}{2}$,

$$T(t) = \begin{cases} T_{A_j} & \text{during antenna half-cycles} \\ T_R & \text{during reference half-cycles} \end{cases} \quad (23)$$

where T_{A_j} is the antenna temperature of the j -th pixel other than the observed pixel. T_{A_j} may or may not equal T_A for the observed pixel.

But since, for $|t-t_o| > \frac{\tau_p}{2}$,

$$C_N(t) = \begin{cases} 0 & \text{during antenna half-cycles} \\ -\frac{1}{N} & \text{during reference half-cycles} \end{cases} \quad (24)$$

the following relation holds:

$$C_N(t) [T(t) + T_{RN}] = (T_R + T_{RN}) C_N(t) \quad (25)$$

Therefore, the remaining portion of the integral in Equation (12) is

$$\begin{aligned} & \frac{2}{\tau_p} \int_{|t-t_o| > \frac{\tau_p}{2}} C_N(t) G(t) [T(t) + T_{RN}] dt \\ &= (T_R + T_{RN}) \frac{2}{\tau_p} \int_{|t-t_o| > \frac{\tau_p}{2}} C_N(t) G(t) dt \end{aligned} \quad (26)$$

Adding (22) and (26), we find that

$$\begin{aligned}
 V_o &= \frac{2}{\tau_p} \int_{N\tau_p} C_N(t)G(t) \left[T(t) + T_{RN} \right] dt \\
 &= \left(\bar{G} + \frac{\gamma_1}{2} \right) (T_A - T_R) + (T_R + T_{RN}) \frac{2}{\tau_p} \int_{N\tau_p} C_N(t)G(t) dt \\
 &= \left(\bar{G} + \frac{\gamma_1}{2} \right) (T_A - T_R) + \gamma_N (T_R + T_{RN})
 \end{aligned} \tag{27}$$

where we define

$$\gamma_N = \frac{2}{\tau_p} \int_{N\tau_p} C_N(t)G(t) dt \tag{28}$$

If we rewrite Equation (17) for the Dicke case as follows:

$$V_o = \left(\bar{G} + \frac{\gamma_1}{2} \right) (T_A - T_R) + \gamma_1 (T_R + T_{KN}) \tag{17}$$

we see that (27) also describes the Dicke case when $\gamma_N = \gamma_1$. The slope of the linear relation between V_o and T_A is the same for the Dicke as for the reference averaging case, but the intercept is different.

Taking the differential of Equation (27), we obtain the following expression for the antenna temperature error due to gain variations in a reference averaging radiometer:

$$\delta T_A = \frac{1}{\bar{G} + \frac{\gamma_1}{2}} \left[\left(\delta \bar{G} + \frac{\delta \gamma_1}{2} \right) (T_A - T_R) + \delta \gamma_N (T_R + T_{RN}) \right] \tag{29}$$

where $\delta\bar{G}$, $\delta\gamma_1$ and $\delta\gamma_N$ are the variations in the gain dependent quantities $\bar{G}$, γ_1 and γ_N from their values at the last calibration. The above formula for $N = 1$ applies to the Dicke radiometer, and is the same as Equation (19).

One should distinguish between the two terms appearing above. The first term is multiplied by $(T_A - T_R)$, and can therefore be suppressed by keeping T_R close to T_A . We may call it a "differential radiometer type" gain variation error. The second term, on the other hand, is multiplied by $T_R + T_{RN}$, the system noise temperature during the reference look. It behaves like a "total power radiometer type" gain variation error, and is a more sensitive term. Moreover, it contains the dependence of δT_A on the parameter N . In particular, the difference between the δT_A for a Dicke and δT_A for reference averaging over N pixels lies entirely in the second term. Therefore, our attention will be focused on analyzing the behavior of the second terms in Equations (27) and (29).

4. Gain Variation Error Dependence on N and Chopping Frequency

Up to this point, our formalism is applicable to arbitrary $G(t)$ functions. For any given model of $G(t)$, we first evaluate $\bar{G}$, γ_1 and γ_N from Equations (15), (16) and (28), then substitute these inputs into Equation (27) to find the V_o response; and then determine the antenna temperature error from Equation (29) by considering the variation of these inputs. This analytical procedure can be readily coded in a computer program, to handle $G(t)$ functions of arbitrary complexity.

In the following analysis, we apply our general formalism to a simple model for $G(t)$, which can describe low frequency gain changes approximately. Our objective is to investigate the dependence of the gain variation error on the parameters N , τ_c and τ_p , and to obtain formulas that can be used to find order-of-magnitude estimates of the increase in gain variation error traded for reduction of detector noise error when reference averaging is used.

We will assume that the variations in $G(t)$ within an interval $N\tau_p$ are slow enough that they can be approximated by a cubic equation of the form

$$G(t) \approx \sum_{n=0}^3 a_n (t - t_o)^n \quad \text{for} \quad |t - t_o| \leq \frac{N\tau_p}{2} \quad (30)$$

where a_n are constants, and t_o is the time at the center of the observed pixel interval, which is also the center of the reference averaging interval. The range of N values for which this approximation is valid will depend on τ_p and on the frequencies of the significant components

in the gain variation spectrum. Should this cubic model be inadequate to cover the interested range of N values for the gain variation frequencies encountered, one could repeat this analysis with higher order models for $G(t)$.

In the analysis to follow, integrals involving $C_N(t)$ will be evaluated, and the following assumptions will be made:

(1) That $\frac{\tau_p}{\tau_c} = n_c$ is an integer, n_c being the number of chop cycles per pixel time. Consequently,

$$\int_{N\tau_p} C_N(t) \times \text{constant } dt = 0 \quad (31)$$

This equality follows from the fact that $C_N(t)$ has n_c positive pulses of height $+1$ and Nn_c negative pulses of height $-\frac{1}{N}$, all the pulses being of equal widths.

(2) That $C_N(t)$ is a symmetric function of $t-t_0$, where t_0 is the center of the observed pixel dwell time and also the center of the integration range. This requires, first of all, symmetric reference averaging, which we have assumed all along. In addition, it requires letting the pixel-to-pixel transition time fall, not at an antenna-reference chopper transition time, but a quarter chop cycle later or before, as shown in Figure 2. The same phasing also applies to the end points of the total integration range. The reference averaging algorithm should take in the appropriate fraction of a chop on each end of the integration range to make $C_N(t)$ symmetric about t_0 . As a consequence of this symmetry,

$$\int_{N\tau_p} C_N(t) \cdot (t - t_o)^n dt = 0 \quad (32)$$

for n equal to an odd integer.

(3) That $\frac{\tau_c}{\tau_p} \ll 1$, which allows us to use certain approximations to evaluate integrals involving $C_N(t)$. In practice, $\frac{\tau_c}{\tau_p}$ may range from $\frac{1}{100}$ to $\frac{1}{10}$, depending on chopper design.

Substituting the $G(t)$ given in (30) into Equation (28),

$$\gamma_N = \frac{2a_2}{\tau_p} \int_{N\tau_p} C_N(t) (t-t_o)^2 dt \quad (33)$$

since the a_0 , a_1 and a_3 terms do not contribute because of (31) and (32). Changing the integration variable to $t-t_o$, and noting the symmetry of the integrand, the above expression may be rewritten as

$$\gamma_N = \frac{4a_2}{\tau_p} \int_0^{N\tau_p/2} C_N(t) t^2 dt \quad (34)$$

For $N = 1$,

$$\begin{aligned} \gamma_1 &= \frac{4a_2}{\tau_p} \int_0^{\tau_p/2} C_1(t) t^2 dt \\ &\approx \frac{4a_2}{\tau_p} \cdot \left(-\frac{\tau_c \tau_p^2}{16} \right) = -\frac{a_2 \tau_c \tau_p}{4} \end{aligned} \quad (35)$$

where the integral has been evaluated approximately by a technique similar to that of Attachment 1. The negative value for the integral assumes that $C_1(t)$ starts out at $t = 0$ with $+1$, which would be the case if there are an odd number of chop cycles per pixel. If there are an even number of chop cycles per pixel, $C_1(t)$ would start out at $t = 0$ with -1 , and the sign of the integral would be positive, but the magnitude remains the same.

In Attachment 1, we defined $p(t)$ as the pulse-train function of period τ_c , which alternately equals $+1$ and 0 every $\frac{\tau_c}{2}$ seconds. $p(t)$ is shown in Figure A-1 of Attachment 1. Identifying the pulse width, $\Delta t = \frac{\tau_c}{2}$, $C_N(t)$ can be expressed in terms of $p(t)$ as follows:

$$C_N(t) = \begin{cases} p(t) - \frac{1}{N} p\left(t - \frac{\tau_c}{2}\right) & 0 < t < \frac{\tau_p}{2} \\ -\frac{1}{N} p\left(t - \frac{\tau_c}{2}\right) & t > \frac{\tau_p}{2} \end{cases} \quad (36)$$

Therefore, the integral that appears in (34) is given by

$$\begin{aligned} \int_0^{\frac{N\tau_p}{2}} C_N(t) t^2 dt &= \int_0^{\frac{\tau_p}{2}} p(t) t^2 dt - \frac{1}{N} \int_0^{\frac{N\tau_p}{2}} p\left(t - \frac{\tau_c}{2}\right) t^2 dt \\ &= \int_0^{\frac{\tau_p}{2}} p(t) t^2 dt - \frac{1}{N} \int_{\frac{\tau_c}{2}}^{\frac{N\tau_p}{2} + \frac{\tau_c}{2}} p(t) t^2 dt \end{aligned} \quad (37)$$

In Attachment 1, we proved the following approximation:

$$\int_{t_a}^{t_b} p(t)t^2 dt \approx \frac{t_b^3 - t_a^3}{6} - \frac{t_b^2 - t_a^2}{4} \left(\frac{\tau_c}{2} \right) \quad (38)$$

By this formula,

$$\int_0^{\frac{\tau_p}{2}} p(t)t^2 dt = \left(\frac{1}{3} - \frac{\tau_c}{2\tau_p} \right) \left(\frac{\tau_p^3}{16} \right) \quad (39)$$

$$\int_{\frac{\tau_c}{2}}^{\frac{N\tau_p}{2} + \frac{\tau_c}{2}} p(t)t^2 dt = \left(\frac{N^3}{3} + \frac{N^2}{2} \cdot \frac{\tau_c}{\tau_p} \right) \left(\frac{\tau_p^3}{16} \right) \quad (40)$$

Substituting these into (37) gives

$$\int_0^{\frac{N\tau_p}{2}} C_N(t)t^2 dt = \left(-\frac{N^2-1}{3} - \frac{N+1}{2} \cdot \frac{\tau_c}{\tau_p} \right) \cdot \frac{\tau_p^3}{16} \quad (41)$$

Equation (34) becomes

$$\gamma_N = -\frac{a_2 \tau_p^2}{4} \left(\frac{N^2-1}{3} + \frac{N+1}{2} \cdot \frac{\tau_c}{\tau_p} \right) \quad (42)$$

Returning to Equation (29), let us distinguish that part of δT_A which depends on N by the symbol $\hat{\delta T}_A$:

$$\begin{aligned}\hat{\delta T}_A &= \frac{1}{\bar{G} + \frac{\gamma_1}{2}} \delta \gamma_N \cdot (T_R + T_{RN}) \\ &\approx \frac{\delta \gamma_N}{\bar{G}} (T_R + T_{RN})\end{aligned}\quad (43)$$

where $\frac{\gamma_1}{2}$ has been neglected in comparison with $\bar{G}$ because

$$\left| \frac{\gamma_1}{2} \right| = \left| \frac{a_2 \tau_p^2}{8} \cdot \frac{\tau_c}{\tau_p} \right| \ll \bar{G} \quad (44)$$

which follows from $\frac{\tau_c}{\tau_p} \ll 1$, and from the expected upper bound on gain variability of, say,

$$\frac{\left| a_2 \left(\frac{\tau_p}{2} \right)^2 \right|}{\bar{G}} < \text{few percent} \quad (45)$$

Combining (42) and (43), we find that $\hat{\delta T}_A$ for reference averaging is given by

$$\hat{\delta T}_A = - \frac{\delta a_2 \tau_p^2}{4\bar{G}} \left(\frac{N^2 - 1}{3} + \frac{N+1}{2} \frac{\tau_c}{\tau_p} \right) (T_R + T_{RN}) \quad (46)$$

which shows explicitly the dependence on the parameters N and $\frac{\tau_c}{\tau_p}$, the latter being inversely proportional to chopping frequency. Compare this with $\hat{\delta T}_A$ for a Dicke radiometer, which is given by

$$\hat{\delta T}_A = - \frac{\delta a_2 \tau_p^2}{4\bar{G}} \left(\frac{\tau_c}{\tau_p} \right) (T_R + T_{RN}) \quad (47)$$

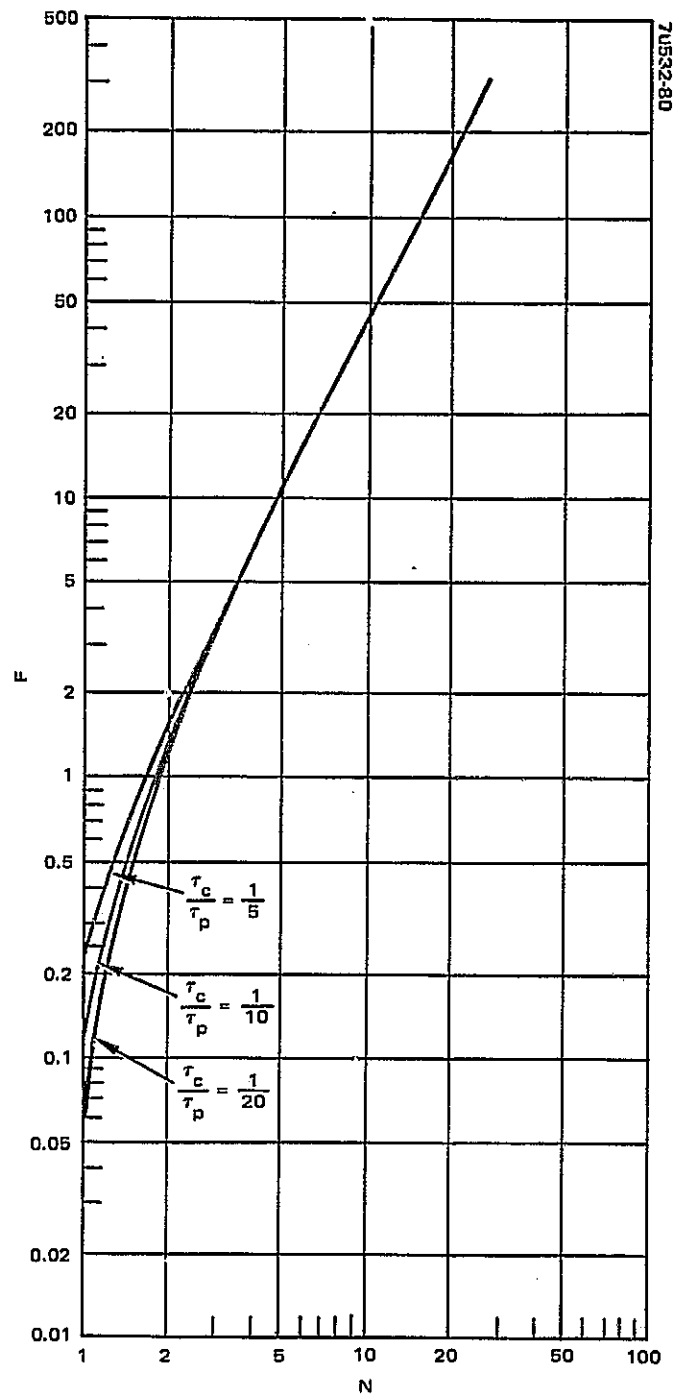


FIGURE 3. PARAMETER F VERSUS N

In the Dicke case, $\delta\hat{T}_A$ is proportional to $\frac{\tau_c}{\tau_p}$, and is reduced to a small value by keeping $\frac{\tau_c}{\tau_p} \ll 1$. In the reference averaging case, $\delta\hat{T}_A$ contains a term independent of $\frac{\tau_c}{\tau_p}$ which rapidly increases with increasing N.

The difference between the two cases lies in the factor appearing in the first parenthesis. We will call this dimensionless factor

$$F = F\left(N, \frac{\tau_c}{\tau_p}\right) = \frac{N^2-1}{3} + \frac{N+1}{2} \frac{\tau_c}{\tau_p} \quad (48)$$

The ratio of $\delta\hat{T}_A$ for reference averaging to $\delta\hat{T}_A$ for a conventional Dicke is given by

$$\frac{F\left(N, \frac{\tau_c}{\tau_p}\right)}{F\left(1, \frac{\tau_c}{\tau_p}\right)} = \frac{\frac{N^2-1}{3} + \frac{N+1}{2} \cdot \frac{\tau_c}{\tau_p}}{\frac{\tau_c}{\tau_p}} \quad (49)$$

For example, if $\frac{\tau_c}{\tau_p} = .1$, going from a conventional Dicke to a reference averaging system with $N=6$ will increase $\delta\hat{T}_A$ by a factor of 120. Hence, while the term $\delta\hat{T}_A$ may be ignored in a Dicke system, it must be dealt with in a reference averaging system.

Figure 3 gives a plot of $F\left(N, \frac{\tau_c}{\tau_p}\right)$.

Note that $\frac{\tau_c}{\tau_p}$ affects F only for low values of N. For $N > 3$, the dependence on $\frac{\tau_c}{\tau_p}$ may be ignored, and

$$F \approx \frac{N^2-1}{3} \quad (50)$$

To give some feeling for the magnitude of $\delta \hat{T}_A$, Equation (46) will be evaluated below for some typical system parameters and gain variabilities. But first, it is useful to express δa_2 in terms of more easily visualized parameters to describe gain variability.

Sometimes gain stability may be specified by such a statement as "gain variations should be less than 1% at .01 Hz". This may be interpreted by a simple model of the form

$$G(t) = G_o + G_w \cos (\omega t + \phi) \quad (51)$$

where G_o is constant, $\frac{G_w}{G_o} < 1\%$, $\frac{\omega}{2\pi} = .01 \text{ Hz}$ and ϕ is random. It can be shown that if

$$N\tau_p < \frac{2}{\omega} \quad (52)$$

$G(t)$ in (51) can be approximated in the interval $-\frac{N\tau_p}{2} < t-t_o < \frac{N\tau_p}{2}$ by the first four terms of its power series expansion about t_o , with errors less than 4% of G_w :

$$G(t) = G(t_o) + G'(t_o)(t-t_o) + \frac{G''(t_o)}{2!}(t-t_o)^2 + \frac{G'''(t_o)}{3!}(t-t_o)^3 \quad (53)$$

Identifying the coefficients with those in Equation (30), we find that

$$a_2 = \frac{G''(t_o)}{2} \quad (54)$$

Now, as ϕ varies randomly; $G''(t_o)$ will vary over the range

$$-\omega^2 G_w < G''(t_o) < \omega^2 G_w$$

Therefore, the worst-case peak-to-peak variation in a_2 is

$$\delta a_2 = \omega^2 G_w \quad (55)$$

Substituting this into (46), and considering only the absolute value of $\delta\hat{T}_A$,

$$\delta\hat{T}_A = \frac{\omega^2 G \tau_p^2}{4\bar{G}} \left(\frac{N^2-1}{3} + \frac{N+1}{2} \frac{\tau_c}{\tau_p} \right) (T_R + T_{RN}) \quad (56)$$

As an example, consider a reference averaging system with the following parameters:

$$T_R + T_{RN} = 3500^\circ K \quad (57)$$

$$\frac{\tau_c}{\tau_p} = \frac{1}{10}$$

$$\tau_p = 1 \text{ sec}$$

and evaluate Equation (56) for gain variabilities described by

$$\frac{\omega}{2\pi} = .01 \text{ Hz}$$

$$\frac{G}{\bar{G}} = .1\%, .5\% \text{ and } 1\%$$

The results are shown in Figure 4 as plots of $\delta\hat{T}_A$ versus N. According to condition (52), these curves are valid for values of N up to

$$N < \frac{100}{\pi} \approx 32.$$

For the case $T_A = T_R$, as in a null balanced radiometer, the first term in Equation (29) vanishes, and the total temperature error due to gain variation is equal to $\delta\hat{T}_A$. This residual error is quite negligible when there is no reference averaging, since its value is $\sim .001^\circ K$, as

indicated by the values of $\delta\hat{T}_A$ for $N = 1$. But when reference averaging is used, this residual error grows rapidly with increasing N , as shown in Figure 4.

For cases where $T_A \neq T_R$, the two terms in Equation (29) must be combined to find the total gain variation error. The manner they are to be combined depends on how $\delta\bar{G}$ and $\delta\gamma_N$ are correlated and on the sign and magnitude of $T_A - T_R$; therefore, for these cases, it is advisable to treat each case separately, according to the specific characteristics of the gain and scene temperature variations, rather than to generalize.

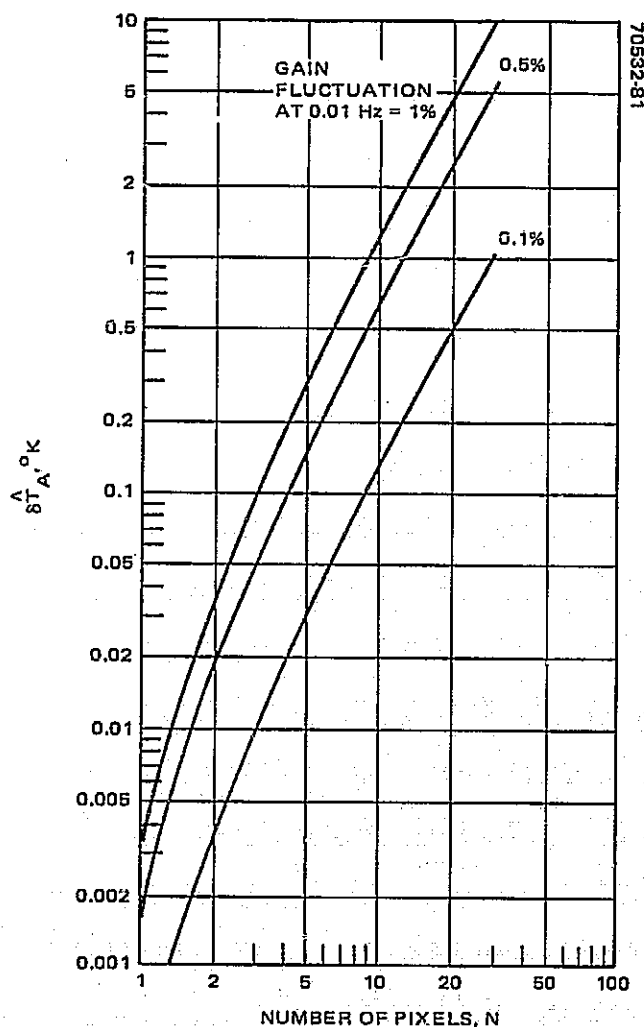


FIGURE 4. ERROR $\delta\hat{T}_A$ VERSUS N

5. Detector Noise vs. Gain Variation Error Tradeoff and Estimates of Maximum Possible Improvement

Since gain variation error increases with N while detector noise error decreases with N , there must be some optimal value of N where the total antenna temperature error is minimum. To illustrate this tradeoff and optimization, we will work out the case for a null-balanced reference averaging radiometer where $T_A = T_R$.

Let δT_{An} and δT_{Ag} denote respectively the detector noise and gain variation errors, which are assumed to be uncorrelated. From Reference 1, δT_{An} is given by

$$\delta T_{An} = \frac{T_{SN}}{\sqrt{B\tau_P}} \sqrt{2 + \frac{2}{N}} \quad (58)$$

where

$$T_{SN} = T_R + T_{RN} = T_A + T_{RN} \quad (59)$$

is the system noise temperature. On the other hand, the gain variation error δT_{Ag} is given by Equation (56); so the total antenna temperature error δT_A consists of

$$\begin{aligned} \delta T_A &= \sqrt{(\delta T_{An})^2 + (\delta T_{Ag})^2} \\ &= \frac{T_{SN}}{\sqrt{B\tau_P}} \left[2 + \frac{2}{N} + \frac{\omega^4 G^2 \omega_{\tau_P}^4}{16G^2} \left(\frac{N^2-1}{3} + \frac{N+1}{2} \frac{\tau_c}{\tau_P} \right)^2 B\tau_P \right]^{\frac{1}{2}} \\ &= \frac{T_{SN}}{\sqrt{B\tau_P}} K \end{aligned} \quad (60)$$

Above, we have generalized the definition of the K-factor to include gain variation error. This generalized K-factor now depends not only on N , but also the dimensionless parameters $\frac{\tau_c}{\tau_p}$, $B\tau_p$, and the gain variability dimensionless parameters $\omega\tau_p$ and $\frac{\omega}{\bar{G}}$.

As an example, we again choose the system parameters given in (57) and assume a gain fluctuation of .1% at .01 Hz. For the MASR water vapor 183 GHz channels, the typical IF bandwidths are 2.5 and 1.25 GHz. With $\tau_p = 1$ sec, we have $B\tau_p = 2.5 \times 10^9$ and 1.25×10^9 , respectively. Assuming the aforementioned system parameters and gain variability, we evaluated the K-factor in Equation (60) as a function of N , for the two values of $B\tau_p$. The results are shown in Figure 5.

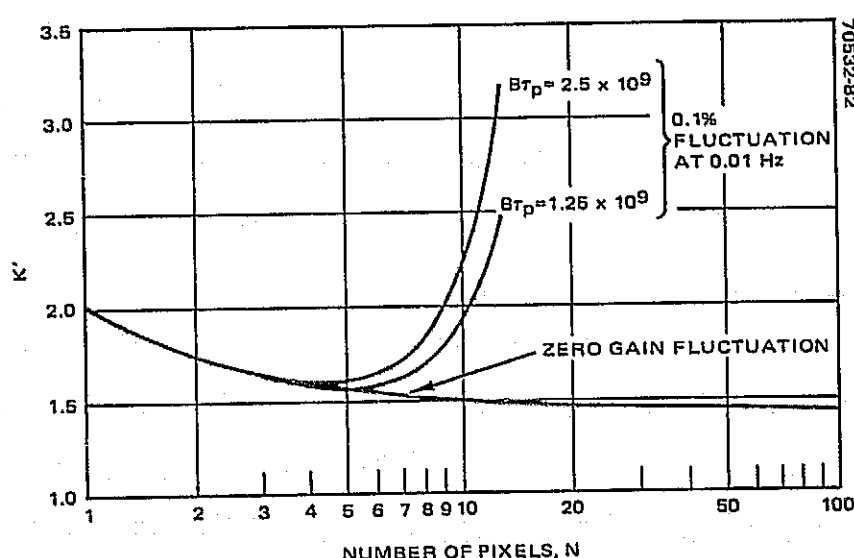


FIGURE 5. GENERALIZED K FACTOR (COMBINED ERROR FROM DETECTOR NOISE AND GAIN FLUCTUATIONS) VERSUS N

ATTACHMENT 1 TO APPENDIX B

The following integral will be evaluated approximately:

$$\int_{t_a}^{t_b} p(t)t^2 dt \quad (A-1)$$

where

$$p(t) = \begin{cases} 1 \\ 0 \end{cases} \quad \text{alternately every } \Delta t \text{ seconds.} \quad (A-2)$$

as illustrated in Figure 2 of Appendix B, and

$$t_b - t_a = 2n\Delta t \quad (A-3)$$

where

$$\Delta t > 0$$

$$n = \text{positive integer} \gg 1 \quad (A-4)$$

The last condition means $\Delta t \ll |t_b - t_a|$.

Let $t_1, t_2, t_3, \dots$ be the time at the centers of the intervals when $p(t) = 1$, as indicated in Figure 2. Consider the approximating sum that enters in the definition of the following integral:

$$\int_{t_a}^{t_b} t^2 dt \approx t_1^2 \Delta t + (t_1 + \Delta t)^2 \Delta t + t_2^2 \Delta t + (t_2 + \Delta t)^2 \Delta t + \dots \quad (A-5)$$

$$\begin{aligned} &\approx 2 \left[t_1^2 \Delta t + t_2^2 \Delta t + \dots \right] \\ &+ \left[2t_1 \Delta t + 2t_2 \Delta t + \dots \right] \Delta t \\ &+ \frac{(\Delta t)^2 (t_b - t_a)}{2} \end{aligned}$$

The sum in the second bracket is approximately equal to the following integral:

$$\int_{t_a - \frac{\Delta t}{2}}^{t_b - \frac{\Delta t}{2}} t dt = \frac{t_b^2 - t_a^2 - (t_b - t_a)\Delta t}{2} \quad (\text{A-6})$$

Whereas, the sum in the first bracket is approximately equal to the integral (A-1), which we are seeking. Therefore, Equation (A-5) is approximately

$$\begin{aligned} \int_{t_a}^{t_b} t^2 dt &\approx 2 \int_{t_a}^{t_b} p(t) t^2 dt + \Delta t \cdot \int_{t_a - \frac{\Delta t}{2}}^{t_b - \frac{\Delta t}{2}} t dt \\ &\quad + \frac{(\Delta t)^2 (t_b - t_a)}{2} \end{aligned} \quad (\text{A-7})$$

Substituting (A-6) into (A-7), we obtain

$$\begin{aligned} 2 \int_{t_a}^{t_b} p(t) t^2 dt &= \int_{t_a}^{t_b} t^2 dt - (\Delta t) \frac{t_b^2 - t_a^2}{2} \\ &= \frac{t_b^3 - t_a^3}{3} - (\Delta t) \frac{t_b^2 - t_a^2}{2} \end{aligned} \quad (\text{A-8})$$

References

1. J. Bremer and R. E. Prince, "Microwave Radiometer Design Study", Operations Research, Inc. Report to Goddard Space Flight Center under Contract NAS 5-23437, Mod 11, January, 1977.
2. R. Graves, "Comments on Reference Averaging in a Dicke Radiometer", rough draft of internal memorandum, August, 1977.
3. D. F. Wait, "The Sensitivity of the Dicke Radiometer", Journal of Research, National Bureau of Standards, Vol. 71C, pp. 127-152 (1967).

APPENDIX C. LASER BEAMS AND RESONATORS

by

Herwig Kogelnik and Tinge Li

Reprinted from

Applied Optics, Vol. 5, No. 10

October 1966

Laser Beams and Resonators

H. KOGELNIK AND T. LI

Abstract—This paper is a review of the theory of laser beams and resonators. It is meant to be tutorial in nature and useful in scope. No attempt is made to be exhaustive in the treatment. Rather, emphasis is placed on formulations and derivations which lead to basic understanding and on results which bear practical significance.

1. INTRODUCTION

THE COHERENT radiation generated by lasers or masers operating in the optical or infrared wavelength regions usually appears as a beam whose transverse extent is large compared to the wavelength. The resonant properties of such a beam in the resonator structure, its propagation characteristics in free space, and its interaction behavior with various optical elements and devices have been studied extensively in recent years. This paper is a review of the theory of laser beams and resonators. Emphasis is placed on formulations and derivations which lead to basic understanding and on results which are of practical value.

Historically, the subject of laser resonators had its origin when Dicke [1], Prokhorov [2], and Schawlow and Townes [3] independently proposed to use the Fabry-Perot interferometer as a laser resonator. The modes in such a structure, as determined by diffraction effects, were first calculated by Fox and Li [4]. Boyd and Gordon [5], and Boyd and Kogelnik [6] developed a theory for resonators with spherical mirrors and approximated the modes by wave beams. The concept of electromagnetic wave beams was also introduced by Goubau and Schwering [7], who investigated the properties of sequences of lenses for the guided transmission of electromagnetic waves. Another treatment of wave beams was given by Pierce [8]. The behavior of Gaussian laser beams as they interact with various optical structures has been analyzed by Goubau [9], Kogelnik [10], [11], and others.

The present paper summarizes the various theories and is divided into three parts. The first part treats the passage of paraxial rays through optical structures and is based on geometrical optics. The second part is an analysis of laser beams and resonators, taking into account the wave nature of the beams but ignoring diffraction effects due to the finite size of the apertures. The third part treats the resonator modes, taking into account aperture diffraction effects. Whenever applicable, useful results are presented in the forms of formulas, tables, charts, and graphs.

Manuscript received July 12, 1966.

H. Kogelnik is with Bell Telephone Laboratories, Inc., Murray Hill, N. J.

T. Li is with Bell Telephone Laboratories, Inc., Holmdel, N. J.

2. PARAXIAL RAY ANALYSIS

A study of the passage of paraxial rays through optical resonators, transmission lines, and similar structures can reveal many important properties of these systems. One such "geometrical" property is the stability of the structure [6], another is the loss of unstable resonators [12]. The propagation of paraxial rays through various optical structures can be described by ray transfer matrices. Knowledge of these matrices is particularly useful as they also describe the propagation of Gaussian beams through these structures; this will be discussed in Section 3. The present section describes briefly some ray concepts which are useful in understanding laser beams and resonators, and lists the ray matrices of several optical systems of interest. A more detailed treatment of ray propagation can be found in textbooks [13] and in the literature on laser resonators [14].

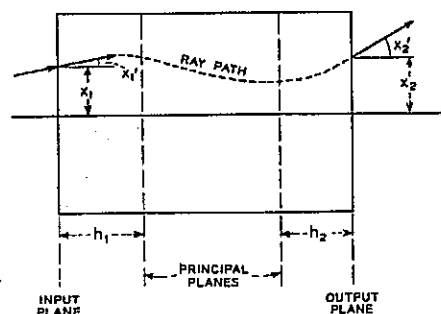


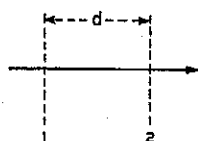
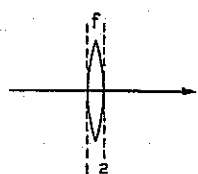
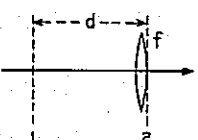
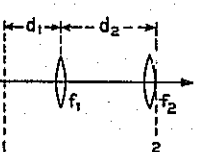
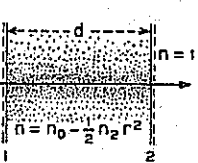
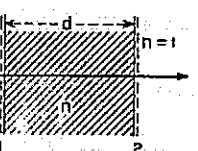
Fig. 1. Reference planes of an optical system. A typical ray path is indicated.

2.1 Ray Transfer Matrix

A paraxial ray in a given cross section ($z = \text{const}$) of an optical system is characterized by its distance x from the optic (z) axis and by its angle or slope x' with respect to that axis. A typical ray path through an optical structure is shown in Fig. 1. The slope x' of paraxial rays is assumed to be small. The ray path through a given structure depends on the optical properties of the structure and on the input conditions, i.e., the position x_1 and the slope x'_1 of the ray in the input plane of the system. For paraxial rays the corresponding output quantities x_2 and x'_2 are linearly dependent on the input quantities. This is conveniently written in the matrix form

$$\begin{bmatrix} x_2 \\ x'_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_1 \\ x'_1 \end{bmatrix} \quad (1)$$

TABLE I
RAY TRANSFER MATRICES OF SIX ELEMENTARY OPTICAL STRUCTURES

NO	OPTICAL SYSTEM	RAY TRANSFER MATRIX
1		$\begin{vmatrix} 1 & d \\ 0 & 1 \end{vmatrix}$
2		$\begin{vmatrix} 1 & 0 \\ -1/f & 1 \end{vmatrix}$
3		$\begin{vmatrix} 1 & d \\ -1/f & 1 - d/f \end{vmatrix}$
4		$\begin{vmatrix} 1 - d_2/f_1 & d_1 + d_2 - d_1 d_2/f_1 \\ -1/f_1 + d_2/f_1 f_2 & 1 - d_1/f_1 - d_2/f_2 + d_1 d_2/f_1 f_2 \end{vmatrix}$
5		$\begin{vmatrix} \cos d \sqrt{n_2/n_0} & \frac{1}{\sqrt{n_0 n_2}} \sin d \sqrt{n_2/n_0} \\ -\sqrt{n_0 n_2} \sin d \sqrt{n_2/n_0} & \cos d \sqrt{n_2/n_0} \end{vmatrix}$
6		$\begin{vmatrix} 1 & d/n \\ 0 & 1 \end{vmatrix}$

where the slopes are measured positive as indicated in the figure. The $ABCD$ matrix is called the ray transfer matrix. Its determinant is generally unity

$$AD - BC = 1. \quad (2)$$

The matrix elements are related to the focal length f of the system and to the location of the principal planes by

$$\begin{aligned} f &= -\frac{1}{C} \\ h_1 &= \frac{D-1}{C} \\ h_2 &= \frac{A-1}{C} \end{aligned} \quad (3)$$

where h_1 and h_2 are the distances of the principal planes from the input and output planes as shown in Fig. 1.

In Table I there are listed the ray transfer matrices of six elementary optical structures. The matrix of No. 1 describes the ray transfer over a distance d . No. 2 describes the transfer of rays through a thin lens of focal length f . Here the input and output planes are immediately to the left and right of the lens. No. 3 is a combination of the first two. It governs rays passing first over a distance d and then through a thin lens. If the sequence is reversed the diagonal elements are interchanged. The matrix of No. 4 describes the rays passing through two structures of the No. 3 type. It is obtained by matrix multiplication. The ray transfer matrix for a lenslike medium of length d is given in No. 5. In this medium the refractive index varies quadratically with the distance r from the optic axis.

$$n = n_0 - \frac{1}{2} n_2 r^2. \quad (4)$$

An index variation of this kind can occur in laser crystals and in gas lenses. The matrix of a dielectric material of index n and length d is given in No. 6. The matrix is referred to the surrounding medium of index 1 and is computed by means of Snell's law. Comparison with No. 1 shows that for paraxial rays the effective distance is shortened by the optically denser material, while, as is well known, the "optical distance" is lengthened.

2.2 Periodic Sequences

Light rays that bounce back and forth between the spherical mirrors of a laser resonator experience a periodic focusing action. The effect on the rays is the same as in a periodic sequence of lenses [15] which can be used as an optical transmission line. A periodic sequence of identical optical systems is schematically indicated in Fig. 2. A single element of the sequence is characterized by its $ABCD$ matrix. The ray transfer through n consecutive elements of the sequence is described by the n th power of this matrix. This can be evaluated by means of Sylvester's theorem

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}^n = \frac{1}{\sin^n \Theta} \quad (5)$$

$$\begin{vmatrix} A \sin n\Theta - \sin(n-1)\Theta & B \sin n\Theta \\ C \sin n\Theta & D \sin n\Theta - \sin(n-1)\Theta \end{vmatrix}$$

where

$$\cos \Theta = \frac{1}{2}(A + D). \quad (6)$$

Periodic sequences can be classified as either *stable* or *unstable*. Sequences are stable when the trace $(A+D)$ obeys the inequality

$$-1 < \frac{1}{2}(A + D) < 1. \quad (7)$$

Inspection of (5) shows that rays passing through a stable sequence are periodically refocused. For unstable systems, the trigonometric functions in that equation become hyperbolic functions, which indicates that the rays become more and more dispersed the further they pass through the sequence.

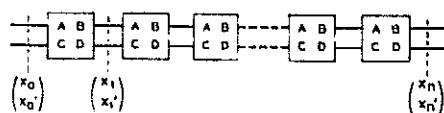


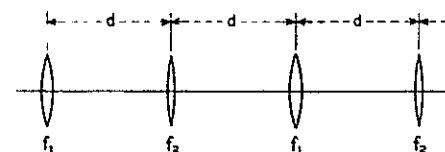
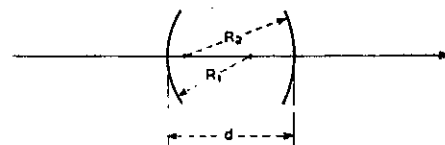
Fig. 2. Periodic sequence of identical systems, each characterized by its $ABCD$ matrix.

2.3 Stability of Laser Resonators

A laser resonator with spherical mirrors of unequal curvature is a typical example of a periodic sequence that can be either stable or unstable [6]. In Fig. 3 such a resonator is shown together with its dual, which is a sequence of lenses. The ray paths through the two structures are the same, except that the ray pattern is folded in the resonator and unfolded in the lens sequence. The focal lengths f_1 and f_2 of the lenses are the same as the focal lengths of the mirrors, i.e., they are determined by the radii of curvature R_1 and R_2 of the mirrors ($f_1 = R_1/2$, $f_2 = R_2/2$). The lens spacings are the same as the mirror spacing d . One can choose, as an element of the periodic sequence, a spacing followed by one lens plus another spacing followed by the second lens. The $ABCD$ matrix of such an element is given in No. 4 of Table I. From this one can obtain the trace, and write the stability condition (7) in the form

$$0 < \left(1 - \frac{d}{R_1}\right) \left(1 - \frac{d}{R_2}\right) < 1. \quad (8)$$

To show graphically which type of resonator is stable and which is unstable, it is useful to plot a stability diagram on which each resonator type is represented by a point. This is shown in Fig. 4 where the parameters d/R_1 and d/R_2 are drawn as the coordinate axes; unstable systems are represented by points in the shaded areas. Various resonator types, as characterized by the relative positions of the centers of curvature of the mirrors, are indicated in the appropriate regions of the diagram. Also entered as alternate coordinate axes are the parameters g_1 and g_2 which play an important role in the diffraction theory of resonators (see Section 4).



$$R_1 = 2f_1, \quad R_2 = 2f_2$$

Fig. 3. Spherical-mirror resonator and the equivalent sequence of lenses.

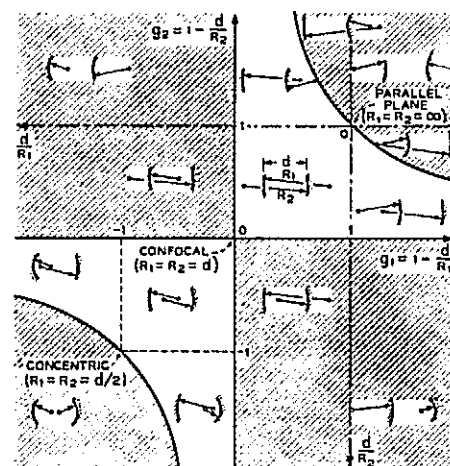


Fig. 4. Stability diagram. Unstable resonator systems lie in shaded regions.

3. WAVE ANALYSIS OF BEAMS AND RESONATORS

In this section the wave nature of laser beams is taken into account, but diffraction effects due to the finite size of apertures are neglected. The latter will be discussed in Section 4. The results derived here are applicable to optical systems with "large apertures," i.e., with apertures that intercept only a negligible portion of the beam power. A theory of light beams or "beam waves" of this kind was first given by Boyd and Gordon [5] and by Goubau and Schwering [7]. The present discussion follows an analysis given in [11].

3.1 Approximate Solution of the Wave Equation

Laser beams are similar in many respects to plane waves; however, their intensity distributions are not uniform, but are concentrated near the axis of propagation and their phase fronts are slightly curved. A field component or potential u of the coherent light satisfies the scalar wave equation

$$\nabla^2 u + k^2 u = 0 \quad (9)$$

where $k = 2\pi/\lambda$ is the propagation constant in the medium.

For light traveling in the z direction one writes

$$u = \psi(x, y, z) \exp(-jkz) \quad (10)$$

where ψ is a slowly varying complex function which represents the differences between a laser beam and a plane wave, namely: a nonuniform intensity distribution, expansion of the beam with distance of propagation, curvature of the phase front, and other differences discussed below. By inserting (10) into (9) one obtains

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} - 2jk \frac{\partial \psi}{\partial z} = 0 \quad (11)$$

where it has been assumed that ψ varies so slowly with z that its second derivative $\partial^2 \psi / \partial z^2$ can be neglected.

The differential equation (11) for ψ has a form similar to the time dependent Schrödinger equation. It is easy to see that

$$\psi = \exp \left\{ -j \left(P + \frac{k}{2q} r^2 \right) \right\} \quad (12)$$

is a solution of (11), where

$$r^2 = x^2 + y^2. \quad (13)$$

The parameter $P(z)$ represents a complex phase shift which is associated with the propagation of the light beam, and $q(z)$ is a complex beam parameter which describes the Gaussian variation in beam intensity with the distance r from the optic axis, as well as the curvature of the phase front which is spherical near the axis. After insertion of (12) into (11) and comparing terms of equal powers in r one obtains the relations

$$q' = 1 \quad (14)$$

and

$$P' = -\frac{j}{q} \quad (15)$$

where the prime indicates differentiation with respect to z . The integration of (14) yields

$$q_2 = q_1 + z \quad (16)$$

which relates the beam parameter q_2 in one plane (output plane) to the parameter q_1 in a second plane (input plane) separated from the first by a distance z .

3.2 Propagation Laws for the Fundamental Mode

A coherent light beam with a Gaussian intensity profile as obtained above is not the only solution of (11), but is perhaps the most important one. This beam is often called the "fundamental mode" as compared to the higher order modes to be discussed later. Because of its importance it is discussed here in greater detail.

For convenience one introduces two real beam parameters R and w related to the complex parameter q by

$$\frac{1}{q} = \frac{1}{R} - j \frac{\lambda}{\pi w^2} \quad (17)$$

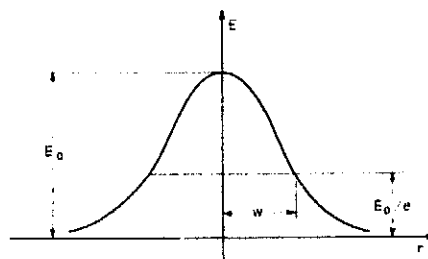


Fig. 5. Amplitude distribution of the fundamental beam.

When (17) is inserted in (12) the physical meaning of these two parameters becomes clear. One sees that $R(z)$ is the radius of curvature of the wavefront that intersects the axis at z , and $w(z)$ is a measure of the decrease of the field amplitude E with the distance from the axis. This decrease is Gaussian in form, as indicated in Fig. 5, and w is the distance at which the amplitude is $1/e$ times that on the axis. Note that the intensity distribution is Gaussian in every beam cross section, and that the width of that Gaussian intensity profile changes along the axis. The parameter w is often called the beam radius or "spot size," and $2w$, the beam diameter.

The Gaussian beam contracts to a minimum diameter $2w_0$ at the beam waist where the phase front is plane. If one measures z from this waist, the expansion laws for the beam assume a simple form. The complex beam parameter at the waist is purely imaginary

$$q_0 = j \frac{\pi w_0^2}{\lambda} \quad (18)$$

and a distance z away from the waist the parameter is

$$q = q_0 + z = j \frac{\pi w_0^2}{\lambda} + z. \quad (19)$$

After combining (19) and (17) one equates the real and imaginary parts to obtain

$$w^2(z) = w_0^2 \left[1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2 \right] \quad (20)$$

and

$$R(z) = z \left[1 + \left(\frac{\pi w_0^2}{\lambda z} \right)^2 \right]. \quad (21)$$

Figure 6 shows the expansion of the beam according to (20). The beam contour $w(z)$ is a hyperbola with asymptotes inclined to the axis at an angle

$$\theta = \frac{\lambda}{\pi w_0^2} \quad (22)$$

This is the far-field diffraction angle of the fundamental mode.

Dividing (21) by (20), one obtains the useful relation

$$\frac{\lambda z}{\pi w_0^2} = \frac{\pi w^2}{\lambda R} \quad (23)$$

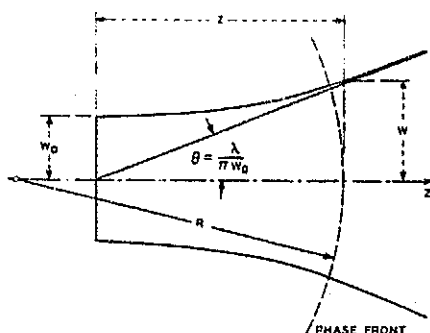


Fig. 6. Contour of a Gaussian beam.

which can be used to express w_0 and z in terms of w and R :

$$w_0^2 = w^2 / \left[1 + \left(\frac{\pi w}{\lambda R} \right)^2 \right] \quad (24)$$

$$z = R / \left[1 + \left(\frac{\lambda R}{\pi w^2} \right)^2 \right] \quad (25)$$

To calculate the complex phase shift a distance z away from the waist, one inserts (19) into (15) to get

$$P' = -\frac{j}{q} = -\frac{j}{z + j(\pi w_0^2/\lambda)} \quad (26)$$

Integration of (26) yields the result

$$\begin{aligned} \int P(z) dz &= \ln[1 - j(\lambda z/\pi w_0^2)] \\ &= \ln \sqrt{1 + (\lambda z/\pi w_0^2)^2} - j \arctan(\lambda z/\pi w_0^2). \end{aligned} \quad (27)$$

The real part of P represents a phase shift difference Φ between the Gaussian beam and an ideal plane wave, while the imaginary part produces an amplitude factor w_0/w which gives the expected intensity decrease on the axis due to the expansion of the beam. With these results for the fundamental Gaussian beam, (10) can be written in the form

$$u(r, z) = \frac{w_0}{w} \exp \left\{ -j(kz - \Phi) - r^2 \left(\frac{1}{w^2} + \frac{jk}{2R} \right) \right\} \quad (28)$$

where

$$\Phi = \arctan(\lambda z/\pi w_0^2). \quad (29)$$

It will be seen in Section 3.5 that Gaussian beams of this kind are produced by many lasers that oscillate in the fundamental mode.

3.3 Higher Order Modes

In the preceding section only one solution of (11) was discussed, i.e., a light beam with the property that its intensity profile in every beam cross section is given by the same function, namely, a Gaussian. The width of this Gaussian distribution changes as the beam propagates along its axis. There are other solutions of (11) with sim-

ilar properties, and they are discussed in this section. These solutions form a complete and orthogonal set of functions and are called the "modes of propagation." Every arbitrary distribution of monochromatic light can be expanded in terms of these modes. Because of space limitations the derivation of these modes can only be sketched here.

a) *Modes in Cartesian Coordinates:* For a system with a rectangular (x, y, z) geometry one can try a solution for (11) of the form

$$\psi = g\left(\frac{x}{w}\right) \cdot h\left(\frac{y}{w}\right) \cdot \exp \left\{ -j \left[P + \frac{k}{2q} (x^2 + y^2) \right] \right\} \quad (30)$$

where g is a function of x and z , and h is a function of y and z . For real g and h this postulates mode beams whose intensity patterns scale according to the width $2w(z)$ of a Gaussian beam. After inserting this trial solution into (11) one arrives at differential equations for g and h of the form

$$\frac{d^2 H_m}{dx^2} - 2x \frac{dH_m}{dx} + 2mH_m = 0. \quad (31)$$

This is the differential equation for the Hermite polynomial $H_m(x)$ of order m . Equation (11) is satisfied if

$$g \cdot h = H_m \left(\sqrt{2} \frac{x}{w} \right) H_n \left(\sqrt{2} \frac{y}{w} \right) \quad (32)$$

where m and n are the (transverse) mode numbers. Note that the same pattern scaling parameter $w(z)$ applies to modes of all orders.

Some Hermite polynomials of low order are

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= x \\ H_2(x) &= 4x^2 - 2 \\ H_3(x) &= 8x^3 - 12x. \end{aligned} \quad (33)$$

Expression (28) can be used as a mathematical description of higher order light beams, if one inserts the product $g \cdot h$ as a factor on the right-hand side. The intensity pattern in a cross section of a higher order beam is, thus, described by the product of Hermite and Gaussian functions. Photographs of such mode patterns are shown in Fig. 7. They were produced as modes of oscillation in a gas laser oscillator [16]. Note that the number of zeros in a mode pattern is equal to the corresponding mode number, and that the area occupied by a mode increases with the mode number.

The parameter $R(z)$ in (28) is the same for all modes, implying that the phase-front curvature is the same and changes in the same way for modes of all orders. The phase shift Φ , however, is a function of the mode numbers. One obtains

$$\Phi(m, n; z) = (m + n + 1) \arctan(\lambda z/\pi w_0^2). \quad (34)$$

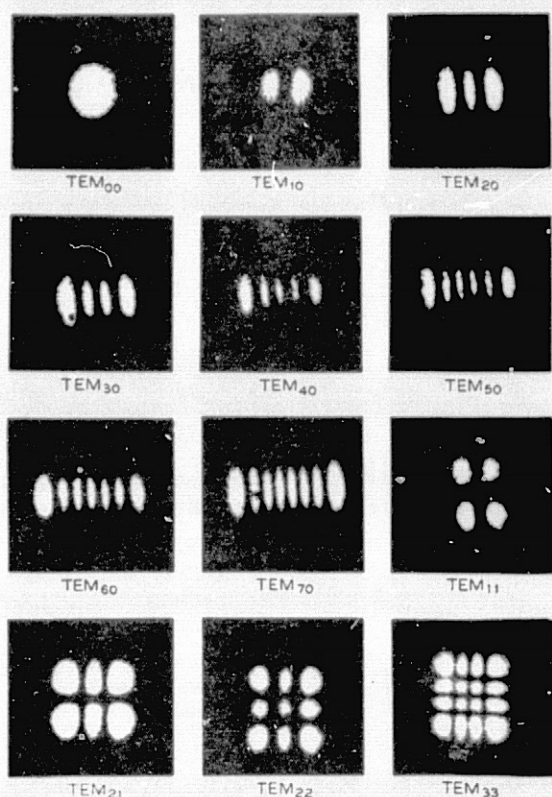


Fig. 7. Mode patterns of a gas laser oscillator (rectangular symmetry).

This means that the phase velocity increases with increasing mode number. In resonators this leads to differences in the resonant frequencies of the various modes of oscillation.

b) *Modes in Cylindrical Coordinates:* For a system with a cylindrical (r, ϕ, z) geometry one uses a trial solution for (11) of the form

$$\psi = g\left(\frac{r}{w}\right) \cdot \exp \left\{ -j \left(P + \frac{k}{2q} r^2 + l\phi \right) \right\}. \quad (35)$$

After some calculation one finds

$$g = \left(\sqrt{2} \frac{r}{w} \right)^l \cdot L_p^l \left(2 \frac{r^2}{w^2} \right) \quad (36)$$

where L_p^l is a generalized Laguerre polynomial, and p and l are the radial and angular mode numbers. $L_p^l(x)$ obeys the differential equation

$$x \frac{d^2 L_p^l}{dx^2} + (l+1-x) \frac{dL_p^l}{dx} + pL_p^l = 0. \quad (37)$$

Some polynomials of low order are

$$\begin{aligned} L_0^l(x) &= 1 \\ L_1^l(x) &= l+1-x \\ L_2^l(x) &= \frac{1}{2}(l+1)(l+2) - (l+2)x + \frac{1}{2}x^2. \end{aligned} \quad (38)$$

As in the case of beams with a rectangular geometry, the beam parameters $w(z)$ and $R(z)$ are the same for all cylindrical modes. The phase shift is, again, dependent on the mode numbers and is given by

$$\Phi(p, l; z) = (2p + l + 1) \arctan(\lambda z / \pi w_0^2). \quad (39)$$

3.4 Beam Transformation by a Lens

A lens can be used to focus a laser beam to a small spot, or to produce a beam of suitable diameter and phase-front curvature for injection into a given optical structure. An ideal lens leaves the transverse field distribution of a beam mode unchanged, i.e., an incoming fundamental Gaussian beam will emerge from the lens as a fundamental beam, and a higher order mode remains a mode of the same order after passing through the lens. However, a lens does change the beam parameters $R(z)$ and $w(z)$. As these two parameters are the same for modes of all orders, the following discussion is valid for all orders; the relationship between the parameters of an incoming beam (labeled here with the index 1) and the parameters of the corresponding outgoing beam (index 2) is studied in detail.

An ideal thin lens of focal length f transforms an incoming spherical wave with a radius R_1 immediately to the left of the lens into a spherical wave with the radius R_2 immediately to the right of it, where

$$\frac{1}{R_2} = \frac{1}{R_1} - \frac{1}{f}. \quad (40)$$

Figure 8 illustrates this situation. The radius of curvature is taken to be positive if the wavefront is convex as viewed from $z = \infty$. The lens transforms the phase fronts of laser beams in exactly the same way as those of spherical waves. As the diameter of a beam is the same immediately to the left and to the right of a *thin* lens, the q -parameters of the incoming and outgoing beams are related by

$$\frac{1}{q_2} = \frac{1}{q_1} - \frac{1}{f}, \quad (41)$$

where the q 's are measured at the lens. If q_1 and q_2 are measured at distances d_1 and d_2 from the lens as indicated in Fig. 9, the relation between them becomes

$$q_2 = \frac{(1 - d_2/f)q_1 + (d_1 + d_2 - d_1 d_2/f)}{-(q_1/f) + (1 - d_1/f)}. \quad (42)$$

This formula is derived using (16) and (41).

More complicated optical structures, such as gas lenses, combinations of lenses, or thick lenses, can be thought of as composed of a series of thin lenses at various spacings. Repeated application of (16) and (41) is, therefore, sufficient to calculate the effect of complicated structures on the propagation of laser beams. If the $ABCD$ matrix for the transfer of paraxial rays through the structure is known, the q parameter of the output beam can be calculated from

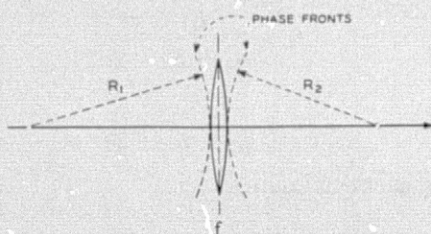


Fig. 8. Transformation of wavefronts by a thin lens.

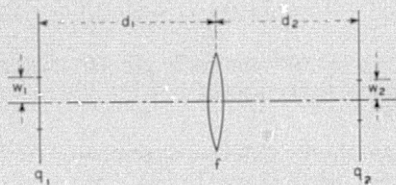


Fig. 9. Distances and parameters for a beam transformed by a thin lens.

$$q_2 = \frac{Aq_1 + R}{Cq_1 + D} \quad (43)$$

This is a generalized form of (42) and has been called the *ABCD* law [10]. The matrices of several optical structures are given in Section II. The *ABCD* law follows from the analogy between the laws for laser beams and the laws obeyed by the spherical waves in geometrical optics. The radius of the spherical waves R obeys laws of the same form as (16) and (41) for the complex beam parameter q . A more detailed discussion of this analogy is given in [11].

3.5 Laser Resonators (Infinite Aperture)

The most commonly used laser resonators are composed of two spherical (or flat) mirrors facing each other. The stability of such "open" resonators has been discussed in Section 2 in terms of paraxial rays. To study the modes of laser resonators one has to take account of their wave nature, and this is done here by studying wave beams of the kind discussed above as they propagate back and forth between the mirrors. As aperture diffraction effects are neglected throughout this section, the present discussion applies only to stable resonators with mirror apertures that are large compared to the spot size of the beams.

A mode of a resonator is defined as a self-consistent field configuration. If a mode can be represented by a wave beam propagating back and forth between the mirrors, the beam parameters must be the same after one complete return trip of the beam. This condition is used to calculate the mode parameters. As the beam that represents a mode travels in both directions between the mirrors it forms the axial standing-wave pattern that is expected for a resonator mode.

A laser resonator with mirrors of equal curvature is shown in Fig. 10 together with the equivalent unfolded system, a sequence of lenses. For this symmetrical structure it is sufficient to postulate self-consistency for one transit of the resonator (which is equivalent to one full period of the lens sequence), instead of a complete return

trip. If the complex beam parameter is given by q_1 , immediately to the right of a particular lens, the beam parameter q_2 , immediately to the right of the next lens, can be calculated by means of (16) and (41) as

$$\frac{1}{q_2} = \frac{1}{q_1 + d} - \frac{1}{f} \quad (44)$$

Self-consistency requires that $q_1 = q_2 = q$, which leads to a quadratic equation for the beam parameter q at the lenses (or at the mirrors of the resonator):

$$\frac{1}{q^2} + \frac{1}{fq} + \frac{1}{fd} = 0 \quad (45)$$

The roots of this equation are

$$\frac{1}{q} = -\frac{1}{2f} (\pm) j \sqrt{\frac{1}{fd} - \frac{1}{4f^2}} \quad (46)$$

where only the root that yields a real beamwidth is used. (Note that one gets a real beamwidth for stable resonators only.)

From (46) one obtains immediately the real beam parameters defined in (17). One sees that R is equal to the radius of curvature of the mirrors, which means that the mirror surfaces are coincident with the phase fronts of the resonator modes. The width $2w$ of the fundamental mode is given by

$$w^2 = \left(\frac{\lambda R}{\pi} \right) / \sqrt{2 \frac{R}{d} - 1} \quad (47)$$

To calculate the beam radius w_0 in the center of the resonator where the phase front is plane, one uses (23) with $z = d/2$ and gets

$$w_0^2 = \frac{\lambda}{2\pi} \sqrt{d(2R - d)} \quad (48)$$

The beam parameters R and w describe the modes of all orders. But the phase velocities are different for the different orders, so that the resonant conditions depend on the mode numbers. Resonance occurs when the phase shift from one mirror to the other is a multiple of π . Using (28) and (34) this condition can be written as

$$kd - 2(m + n + 1) \arctan(\lambda d / 2\pi w_0^2) = \pi(q + 1) \quad (49)$$

where q is the number of nodes of the axial standing-wave pattern (the number of half wavelengths is $q + 1$),¹ and m and n are the rectangular mode numbers defined in Section 3.3. For the modes of circular geometry one obtains a similar condition where $(2p + l + 1)$ replaces $(m + n + 1)$.

The fundamental beat frequency ν_0 , i.e., the frequency spacing between successive longitudinal resonances, is given by

$$\nu_0 = c/2d \quad (50)$$

¹ This q is not to be confused with the complex beam parameter.

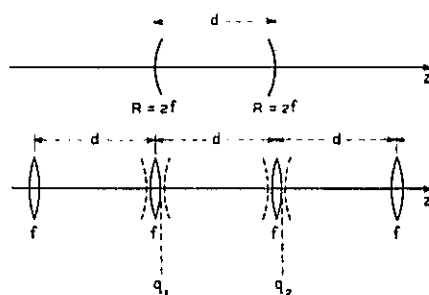


Fig. 10. Symmetrical laser resonator and the equivalent sequence of lenses. The beam parameters, q_1 and q_2 , are indicated.

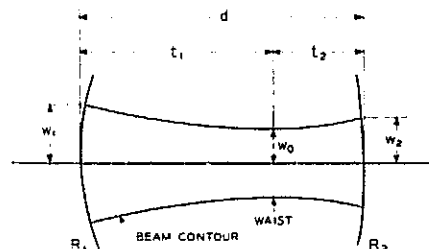


Fig. 11. Mode parameters of interest for a resonator with mirrors of unequal curvature.

where c is the velocity of light. After some algebraic manipulations one obtains from (49) the following formula for the resonant frequency ν of a mode

$$\nu/\nu_0 = (q+1) + \frac{1}{\pi} (m+n+1) \arccos(1-d/R). \quad (51)$$

For the special case of the confocal resonator ($d = R = b$), the above relations become

$$w^2 = \lambda b/\pi, \quad w_0^2 = \lambda b/2\pi; \\ \nu/\nu_0 = (q+1) + \frac{1}{2}(m+n+1). \quad (52)$$

The parameter b is known as the confocal parameter.

Resonators with mirrors of unequal curvature can be treated in a similar manner. The geometry of such a resonator where the radii of curvature of the mirrors are R_1 and R_2 is shown in Fig. 11. The diameters of the beam at the mirrors of a stable resonator, $2w_1$ and $2w_2$, are given by

$$w_1^4 = (\lambda R_1/\pi)^2 \frac{R_2 - d}{R_1 - d} \frac{d}{R_1 + R_2 - d} \\ w_2^4 = (\lambda R_2/\pi)^2 \frac{R_1 - d}{R_2 - d} \frac{d}{R_1 + R_2 - d}. \quad (53)$$

The diameter of the beam waist $2w_0$, which is formed either inside or outside the resonator, is given by

$$w_0^4 = \left(\frac{\lambda}{\pi}\right)^2 \frac{d(R_1 - d)(R_2 - d)(R_1 + R_2 - d)}{(R_1 + R_2 - 2d)^3}. \quad (54)$$

The distances t_1 and t_2 between the waist and the mirrors, measured positive as shown in the figure, are

$$t_1 = \frac{d(R_2 - d)}{R_1 + R_2 - 2d} \\ t_2 = \frac{d(R_1 - d)}{R_1 + R_2 - 2d}. \quad (55)$$

The resonant condition is

$$\nu/\nu_0 = (q+1) + \frac{1}{\pi} (m+n+1) \arccos \sqrt{(1-d/R_1)(1-d/R_2)} \quad (56)$$

where the square root should be given the sign of $(1-d/R_1)$, which is equal to the sign of $(1-d/R_2)$ for a stable resonator.

There are more complicated resonator structures than the ones discussed above. In particular, one can insert a lens or several lenses between the mirrors. But in every case, the unfolded resonator is equivalent to a periodic sequence of identical optical systems as shown in Fig. 2. The elements of the $ABCD$ matrix of this system can be used to calculate the mode parameters of the resonator. One uses the $ABCD$ law (43) and postulates self-consistency by putting $q_1 = q_2 = q$. The roots of the resulting quadratic equation are

$$\frac{1}{q} = \frac{D-A}{2B} \pm \frac{j}{2B} \sqrt{4-(A+D)^2}, \quad (57)$$

which yields, for the corresponding beam radius w ,

$$w^2 = (2\lambda B/\pi) / \sqrt{4-(A+D)^2}. \quad (58)$$

3.6 Mode Matching

It was shown in the preceding section that the modes of laser resonators can be characterized by light beams with certain properties and parameters which are defined by the resonator geometry. These beams are often injected into other optical structures with different sets of beam parameters. These optical structures can assume various physical forms, such as resonators used in scanning Fabry-Perot interferometers or regenerative amplifiers, sequences of dielectric or gas lenses used as optical transmission lines, or crystals of nonlinear dielectric material employed in parametric optics experiments. To match the modes of one structure to those of another one must transform a given Gaussian beam (or higher order mode) into another beam with prescribed properties. This transformation is usually accomplished with a thin lens, but other more complex optical systems can be used. Although the present discussion is devoted to the simple case of the thin lens, it is also applicable to more complex systems, provided one measures the distances from the principal planes and uses the combined focal length f of the more complex system.

The location of the waists of the two beams to be transformed into each other and the beam diameters at the waists are usually known or can be computed. To match the beams one has to choose a lens of a focal length

f that is larger than a characteristic length f_0 defined by the two beams, and one has to adjust the distances between the lens and the two beam waists according to rules derived below.

In Fig. 9 the two beam waists are assumed to be located at distances d_1 and d_2 from the lens. The complex beam parameters at the waists are purely imaginary; they are

$$q_1 = j\pi w_1^2/\lambda, \quad q_2 = j\pi w_2^2/\lambda \quad (59)$$

where $2w_1$ and $2w_2$ are the diameters of the two beams at their waists. If one inserts these expressions for q_1 and q_2 into (42) and equates the imaginary parts, one obtains

$$\frac{d_1 - f}{d_2 - f} = \frac{w_1^2}{w_2^2} \quad (60)$$

Equating the real parts results in

$$(d_1 - f)(d_2 - f) = f^2 - f_0^2 \quad (61)$$

where

$$f_0 = \pi w_1 w_2 / \lambda. \quad (62)$$

Note that the characteristic length f_0 is defined by the waist diameters of the beams to be matched. Except for the term f_0^2 , which goes to zero for infinitely small wavelengths, (61) resembles Newton's imaging formula of geometrical optics.

Any lens with a focal length $f > f_0$ can be used to perform the matching transformation. Once f is chosen, the distances d_1 and d_2 have to be adjusted to satisfy the matching formulas [10]

$$\begin{aligned} d_1 &= f \pm \frac{w_1}{w_2} \sqrt{f^2 - f_0^2}, \\ d_2 &= f \pm \frac{w_2}{w_1} \sqrt{f^2 - f_0^2}. \end{aligned} \quad (63)$$

These relations are derived by combining (60) and (61). In (63) one can choose either both plus signs or both minus signs for matching.

It is often useful to introduce the confocal parameters b_1 and b_2 into the matching formulas. They are defined by the waist diameters of the two systems to be matched

$$b_1 = 2\pi w_1^2/\lambda, \quad b_2 = 2\pi w_2^2/\lambda. \quad (64)$$

Using these parameters one gets for the characteristic length f_0

$$f_0^2 = \frac{1}{4} b_1 b_2, \quad (65)$$

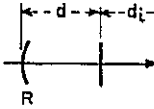
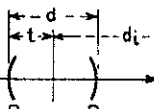
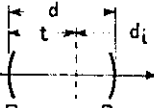
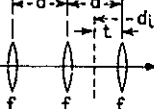
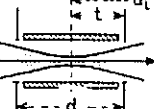
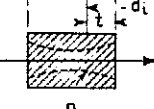
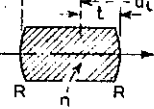
and for the matching distances

$$\begin{aligned} d_1 &= f \pm \frac{1}{2} b_1 \sqrt{(f^2/f_0^2) - 1}, \\ d_2 &= f \pm \frac{1}{2} b_2 \sqrt{(f^2/f_0^2) - 1}. \end{aligned} \quad (66)$$

Note that in this form of the matching formulas, the wavelength does not appear explicitly.

Table II lists, for quick reference, formulas for the two important parameters of beams that emerge from various

TABLE II
FORMULAS FOR THE CONFOCAL PARAMETER AND THE LOCATION OF BEAM WAIST FOR VARIOUS OPTICAL STRUCTURES

NO	OPTICAL SYSTEM	$\frac{1}{2}b = \pi w_0^2/\lambda$	t
1		$\sqrt{d(R-d)}$	-
2		$\frac{1}{2} \sqrt{d(2R-d)}$	$\frac{1}{2} d$
3		$\frac{\sqrt{d(R_1-d)(R_2-d)(R_1+R_2-d)}}{R_1+R_2-2d}$	$\frac{d(R_2-d)}{R_1+R_2-2d}$
4		$\frac{R \sqrt{d(2R-d)}}{2R+d(n^2-1)}$	$\frac{ndR}{2R+d(n^2-1)}$
5		$\frac{1}{2} \sqrt{d(4f-d)}$	$\frac{1}{2} d$
6		$\frac{1}{2} d$	$\frac{1}{2} d$
7		$\frac{d}{2n}$	$\frac{d}{2n}$
8		$\frac{nR \sqrt{d(2R-d)}}{2n^2R-d(n^2-1)}$	$\frac{dR}{2n^2R-d(n^2-1)}$

optical structures commonly encountered. They are the confocal parameter b and the distance t which gives the waist location of the emerging beam. System No. 1 is a resonator formed by a flat mirror and a spherical mirror of radius R . System No. 2 is a resonator formed by two equal spherical mirrors. System No. 3 is a resonator formed by mirrors of unequal curvature. System No. 4

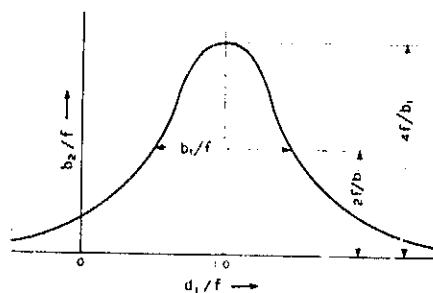


Fig. 12. The confocal parameter b_2 as a function of the lens-waist spacing d_1 .

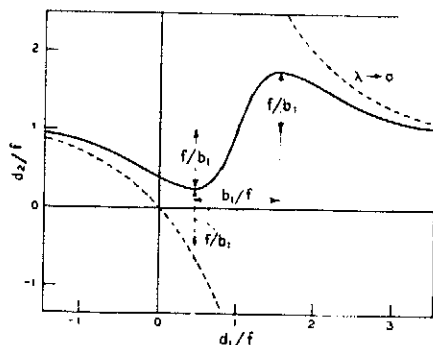


Fig. 13. The waist spacing d_2 as a function of the lens-waist spacing d_1 .

is, again, a resonator formed by two equal spherical mirrors, but with the reflecting surfaces deposited on plano-concave optical plates of index n . These plates act as negative lenses and change the characteristics of the emerging beam. This lens effect is assumed not present in Systems Nos. 2 and 3. System No. 5 is a sequence of thin lenses of equal focal lengths f . System No. 6 is a system of two irises with equal apertures spaced at a distance d . Shown are the parameters of a beam that will pass through both irises with the least possible beam diameter. This is a beam which is "confocal" over the distance d . This beam will also pass through a tube of length d with the optimum clearance. (The tube is also indicated in the figure.) A similar situation is shown in System No. 7, which corresponds to a beam that is confocal over the length d of optical material of index n . System No. 8 is a spherical mirror resonator filled with material of index n , or an optical material with curved end surfaces where the beam passing through it is assumed to have phase fronts that coincide with these surfaces.

When one designs a matching system, it is useful to know the accuracy required of the distance adjustments. The discussion below indicates how the parameters b_2 and d_2 change when b_1 and f are fixed and the lens spacing d_1 to the waist of the input beam is varied. Equations (60) and (61) can be solved for b_2 with the result [9]

$$b_2 f = \frac{b_1 f}{(1 - d_1/f)^2 + (b_1/2f)^2} \quad (67)$$

This means that the parameter b_2 of the beam emerging from the lens changes with d_1 according to a Lorentzian functional form as shown in Fig. 12. The Lorentzian is centered at $d_1=f$ and has a width of b_1 . The maximum value of b_2 is $4f^2/b_1$.

If one inserts (67) into (60) one gets

$$1 - d_2/f = \frac{1 - d_1/f}{(1 - d_1/f)^2 + (b_1/2f)^2} \quad (68)$$

which shows the change of d_2 with d_1 . The change is reminiscent of a dispersion curve associated with a Lorentzian as shown in Fig. 13. The extrema of this curve occur at the halfpower points of the Lorentzian. The slope of the curve at $d_1=f$ is $(2f/b)^2$. The dashed curves in the figure correspond to the geometrical optics imaging relation between d_1 , d_2 , and f [20].

3.7 Circle Diagrams

The propagation of Gaussian laser beams can be represented graphically on a circle diagram. On such a diagram one can follow a beam as it propagates in free space or passes through lenses, thereby affording a graphic solution of the mode matching problem. The circle diagrams for beams are similar to the impedance charts, such as the Smith chart. In fact there is a close analogy between transmission-line and laser-beam problems, and there are analog electric networks for every optical system [17].

The first circle diagram for beams was proposed by Collins [18]. A dual chart was discussed in [19]. The basis for the derivation of these charts are the beam propagation laws discussed in Section 3.2. One combines (17) and (19) and eliminates q to obtain

$$\left(\frac{\lambda}{\pi w^2} + j \frac{1}{R}\right) \left(\frac{\pi w_0^2}{\lambda} - jz\right) = 1. \quad (69)$$

This relation contains the four quantities w , R , w_0 , and z which were used to describe the propagation of Gaussian beams in Section 3.2. Each pair of these quantities can be expressed in complex variables W and

$$W = \frac{\lambda}{\pi} + j \frac{1}{R}$$

$$Z = \frac{\pi w_0^2}{\lambda} - jz = b/2 - jz, \quad (70)$$

where b is the confocal parameter of the beam. For these variables (69) defines a conformal transformation

$$W = 1/Z. \quad (71)$$

The two dual circle diagrams are plotted in the complex planes of W and Z , respectively. The W -plane diagram [18] is shown in Fig. 14 where the variables $\lambda/\pi w^2$ and $1/R$ are plotted as axes. In this plane the lines of constant $b/2 = \pi w_0^2/\lambda$ and the lines of constant z of the Z plane appear as circles through the origin. A beam is represented by a circle of constant b , and the beam parameters w and R at a distance z from the beam waist can be easily read

ORIGINAL PAGE IS
OF POOR QUALITY

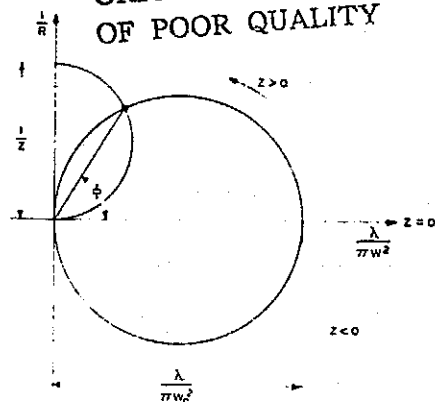


Fig. 14. Geometry for the W -plane circle diagram.

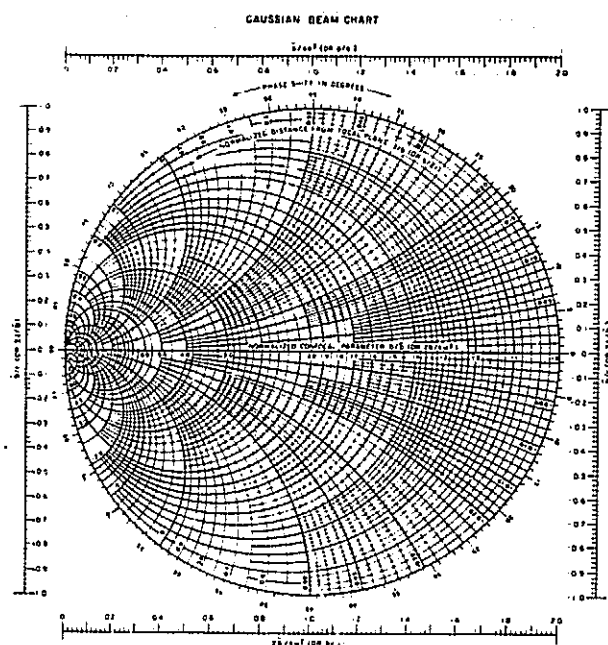


Fig. 15. The Gaussian beam chart. Both W -plane and Z -plane circle diagram are combined into one.

from the diagram. When the beam passes through a lens the phase front is changed according to (40) and a new beam is formed, which implies that the incoming and outgoing beams are connected in the diagram by a vertical line of length $1/f$. The angle ϕ shown in the figure is equal to the phase shift experienced by the beam as given by (29): this is easily shown using (23).

The dual diagram [19] is plotted in the Z plane. The sets of circles in both diagrams have the same form, and only the labeling of the axes and circles is different. In Fig. 15 both diagrams are unified in one chart. The labels in parentheses correspond to the Z -plane diagram, and b is a normalizing parameter which can be arbitrarily chosen for convenience.

One can plot various other circle diagrams which are related to the above by conformal transformations. One

such transformation makes it possible to use the Smith chart for determining complex mismatch coefficients for Gaussian beams [20]. Other circle diagrams include those for optical resonators [21] which allow the graphic determination of certain parameters of the resonator modes.

4. LASER RESONATORS (FINITE APERTURE)

4.1 General Mathematical Formulation

In this section aperture diffraction effects due to the finite size of the mirrors are taken into account; these effects were neglected in the preceding sections. There, it was mentioned that resonators used in laser oscillators usually take the form of an open structure consisting of a pair of mirrors facing each other. Such a structure with finite mirror apertures is intrinsically lossy and, unless energy is supplied to it continuously, the electromagnetic field in it will decay. In this case a mode of the resonator is a *slowly decaying* field configuration whose relative distribution does not change with time [4]. In a laser oscillator the active medium supplies enough energy to overcome the losses so that a steady-state field can exist. However, because of nonlinear gain saturation the medium will exhibit less gain in those regions where the field is high than in those where the field is low, and so the oscillating modes of an active resonator are expected to be somewhat different from the decaying modes of the passive resonator. The problem of an active resonator filled with a saturable-gain medium has been solved recently [22], [23], and the computed results show that if the gain is not too large the resonator modes are essentially unperturbed by saturation effects. This is fortunate as the results which have been obtained for the passive resonator can also be used to describe the active modes of laser oscillators.

The problem of the open resonator is a difficult one and a rigorous solution is yet to be found. However, if certain simplifying assumptions are made, the problem becomes tractable and physically meaningful results can be obtained. The simplifying assumptions involve essentially the quasi-optic nature of the problem; specifically, they are 1) that the dimensions of the resonator are large compared to the wavelength and 2) that the field in the resonator is substantially transverse electromagnetic (TEM). So long as those assumptions are valid, the Fresnel-Kirchhoff formulation of Huygens' principle can be invoked to obtain a pair of integral equations which relate the fields of the two opposing mirrors. Furthermore, if the mirror separation is large compared to mirror dimensions and if the mirrors are only slightly curved, the two orthogonal Cartesian components of the vector field are essentially uncoupled, so that separate scalar equations can be written for each component. The solutions of these scalar equations yield resonator modes which are uniformly polarized in one direction. Other polarization configurations can be constructed from the uniformly polarized modes by linear superposition.

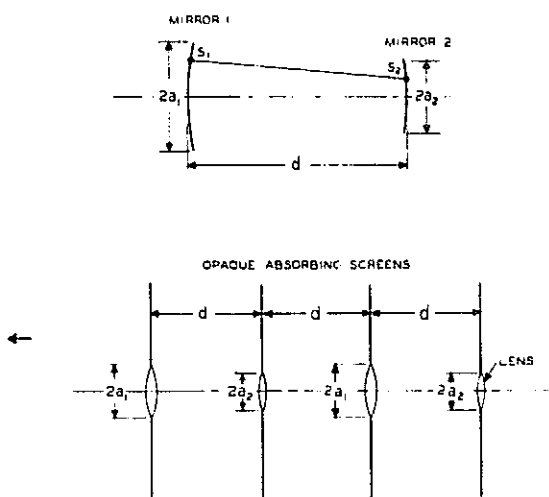


Fig. 16. Geometry of a spherical-mirror resonator with finite mirror apertures and the equivalent sequence of lenses set in opaque absorbing screens.

In deriving the integral equations, it is assumed that a traveling TEM wave is reflected back and forth between the mirrors. The resonator is thus analogous to a transmission medium consisting of apertures or lenses set in opaque absorbing screens (see Fig. 16). The fields at the two mirrors are related by the equations [24]

$$\begin{aligned}\gamma^{(1)} E^{(1)}(s_1) &= \int_{S_2} K^{(2)}(s_1, s_2) E^{(2)}(s_2) dS_2 \\ \gamma^{(2)} E^{(2)}(s_2) &= \int_{S_1} K^{(1)}(s_2, s_1) E^{(1)}(s_1) dS_1\end{aligned}\quad (72)$$

where the integrations are taken over the mirror surfaces S_2 and S_1 , respectively. In the above equations the subscripts and superscripts one and two denote mirrors one and two; s_1 and s_2 are symbolic notations for transverse coordinates on the mirror surface, e.g., $s_1 = (x_1, y_1)$ and $s_2 = (x_2, y_2)$ or $s_1 = (r_1, \phi_1)$ and $s_2 = (r_2, \phi_2)$; $E^{(1)}$ and $E^{(2)}$ are the relative field distribution functions over the mirrors; $\gamma^{(1)}$ and $\gamma^{(2)}$ give the attenuation and phase shift suffered by the wave in transit from one mirror to the other; the kernels $K^{(1)}$ and $K^{(2)}$ are functions of the distance between s_1 and s_2 and, therefore, depend on the mirror geometry; they are equal [$K^{(1)}(s_2, s_1) = K^{(2)}(s_1, s_2)$] but, in general, are not symmetric [$K^{(1)}(s_2, s_1) \neq K^{(1)}(s_1, s_2)$, $K^{(2)}(s_1, s_2) \neq K^{(2)}(s_2, s_1)$].

The integral equations given by (72) express the field at each mirror in terms of the reflected field at the other; that is, they are single-transit equations. By substituting one into the other, one obtains the double-transit or round-trip equations, which state that the field at each mirror must reproduce itself after a round trip. Since the kernel for each of the double-transit equations is symmetric [24], it follows [25] that the field distribution functions corresponding to the different mode orders are orthogonal over their respective mirror surfaces; that is

$$\begin{aligned}\int_{S_1} E_m^{(1)}(s_1) E_n^{(1)}(s_1) dS_1 &= 0, \quad m \neq n \\ \int_{S_2} E_m^{(2)}(s_2) E_n^{(2)}(s_2) dS_2 &= 0, \quad m \neq n\end{aligned}\quad (73)$$

where m and n denote different mode orders. It is to be noted that the orthogonality relation is non-Hermitian and is the one that is generally applicable to lossy systems.

4.2 Existence of Solutions

The question of the existence of solutions to the resonator integral equations has been the subject of investigation by several authors [26]–[28]. They have given rigorous proofs of the existence of eigenvalues and eigenfunctions for kernels which belong to resonator geometries commonly encountered, such as those with parallel-plane and spherically curved mirrors.

4.3 Integral Equations for Resonators with Spherical Mirrors

When the mirrors are spherical and have rectangular or circular apertures, the two-dimensional integral equations can be separated and reduced to one-dimensional equations which are amenable to solution by either analytical or numerical methods. Thus, in the case of rectangular mirrors [4]–[6], [24], [29], [30], the one-dimensional equations in Cartesian coordinates are the same as those for infinite-strip mirrors; for the x coordinate, they are

$$\begin{aligned}\gamma_x^{(1)} u^{(1)}(x_1) &= \int_{-a_2}^{a_2} K(x_1, x_2) u^{(2)}(x_2) dx_2 \\ \gamma_x^{(2)} u^{(2)}(x_2) &= \int_{-a_1}^{a_1} K(x_1, x_2) u^{(1)}(x_1) dx_1\end{aligned}\quad (74)$$

where the kernel K is given by

$$K(x_1, x_2) = \sqrt{\frac{j}{\lambda d}} \exp \left\{ -\frac{jk}{2d} (g_1 x_1^2 + g_2 x_2^2 - 2x_1 x_2) \right\} \quad (75)$$

Similar equations can be written for the y coordinate, so that $E(x, y) = u(x)v(y)$ and $\gamma = \gamma_x \gamma_y$. In the above equation a_1 and a_2 are the half-widths of the mirrors in the x direction, d is the mirror spacing, k is $2\pi/\lambda$, and λ is the wavelength. The radii of curvature of the mirrors R_1 and R_2 are contained in the factors

$$\begin{aligned}g_1 &= 1 - \frac{d}{R_1} \\ g_2 &= 1 - \frac{d}{R_2}\end{aligned}\quad (76)$$

For the case of circular mirrors [4], [31], [32] the equations are reduced to the one-dimensional form by using

cylindrical coordinates and by assuming a sinusoidal azimuthal variation of the field; that is, $E(r, \phi) = R_l(r)e^{-j\ell\phi}$. The radial distribution functions $R_l^{(1)}$ and $R_l^{(2)}$ satisfy the one-dimensional integral equations:

$$\begin{aligned}\gamma_l^{(1)} R_l^{(1)}(r_1) \sqrt{r_1} &= \int_0^{a_2} K_l(r_1, r_2) R_l^{(2)}(r_2) \sqrt{r_2} dr_2 \\ \gamma_l^{(2)} R_l^{(2)}(r_2) \sqrt{r_2} &= \int_0^{a_1} K_l(r_1, r_2) R_l^{(1)}(r_1) \sqrt{r_1} dr_1\end{aligned}\quad (77)$$

where the kernel K_l is given by

$$\begin{aligned}K_l(r_1, r_2) &= \frac{j^{l+1}}{d} J_l\left(k \frac{r_1 r_2}{d}\right) \sqrt{r_1 r_2} \\ &\cdot \exp\left\{-\frac{j\ell}{2d}(g_1 r_1^2 + g_2 r_2^2)\right\}\end{aligned}\quad (78)$$

and J_l is a Bessel function of the first kind and l th order. In (77), a_1 and a_2 are the radii of the mirror apertures and d is the mirror spacing; the factors g_1 and g_2 are given by (76).

Except for the special case of the confocal resonator [5] ($g_1 = g_2 = 0$), no exact analytical solution has been found for either (74) or (77), but approximate methods and numerical techniques have been employed with success for their solutions. Before presenting results, it is appropriate to discuss two important properties which apply in general to resonators with spherical mirrors; these are the properties of "equivalence" and "stability."

4.4 Equivalent Resonator Systems

The equivalence properties [24], [33] of spherical-mirror resonators are obtained by simple algebraic manipulations of the integral equations. First, it is obvious that the mirrors can be interchanged without affecting the results; that is, the subscripts and superscripts one and two can be interchanged. Second, the diffraction loss and the intensity pattern of the mode remain invariant if both g_1 and g_2 are reversed in sign: the eigenfunctions E and the eigenvalues γ merely take on complex conjugate values. An example of such equivalent systems is that of parallel-plane ($g_1 = g_2 = 1$) and concentric ($g_1 = g_2 = -1$) resonator systems.

The third equivalence property involves the Fresnel number N and the stability factors G_1 and G_2 , where

$$\begin{aligned}N &= \frac{a_1 a_2}{\lambda d} \\ G_1 &= g_1 \frac{a_1}{a_2} \\ G_2 &= g_2 \frac{a_2}{a_1}\end{aligned}\quad (79)$$

If these three parameters are the same for any two resonators, then they would have the same diffraction loss, the

same resonant frequency, and mode patterns that are scaled versions of each other. Thus, the equivalence relations reduce greatly the number of calculations which are necessary for obtaining the solutions for the various resonator geometries.

4.5 Stability Condition and Diagram

Stability of optical resonators has been discussed in Section 2 in terms of geometrical optics. The stability condition is given by (8). In terms of the stability factors G_1 and G_2 , it is

$$0 < G_1 G_2 < 1$$

or

$$0 < g_1 g_2 < 1. \quad (80)$$

Resonators are stable if this condition is satisfied and unstable otherwise.

A stability diagram [6], [24] for the various resonator geometries is shown in Fig. 4 where g_1 and g_2 are the coordinate axes and each point on the diagram represents a particular resonator geometry. The boundaries between stable and unstable (shaded) regions are determined by (80), which is based on geometrical optics. The fields of the modes in stable resonators are more concentrated near the resonator axes than those in unstable resonators and, therefore, the diffraction losses of unstable resonators are much higher than those of stable resonators. The transition, which occurs near the boundaries, is gradual for resonators with small Fresnel numbers and more abrupt for those with large Fresnel numbers. The origin of the diagram represents the confocal system with mirrors of equal curvature ($R_1 = R_2 = d$) and is a point of lowest diffraction loss for a given Fresnel number. The fact that a system with minor deviations from the ideal confocal system may become unstable should be borne in mind when designing laser resonators.

4.6 Modes of the Resonator

The transverse field distributions of the resonator modes are given by the eigenfunctions of the integral equations. As yet, no exact analytical solution has been found for the general case of arbitrary G_1 and G_2 , but approximate analytical expressions have been obtained to describe the fields in stable spherical-mirror resonators [5], [6]. These approximate eigenfunctions are the same as those of the optical beam modes which are discussed in Section 2; that is, the field distributions are given approximately by Hermite-Gaussian functions for rectangular mirrors [5], [6], [34], and by Laguerre-Gaussian functions for circular mirrors [6], [7]. The designation of the resonator modes is given in Section 3.5. (The modes are designated as TEM_{mnq} for rectangular mirrors and TEM_{plq} for circular mirrors.) Figure 7 shows photographs of some of the rectangular mode patterns of a

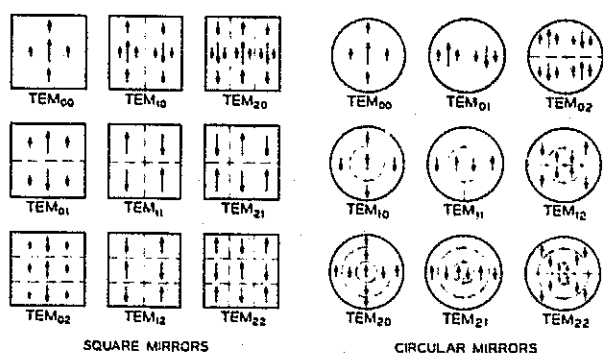


Fig. 17. Linearly polarized resonator mode configurations for square and circular mirrors.

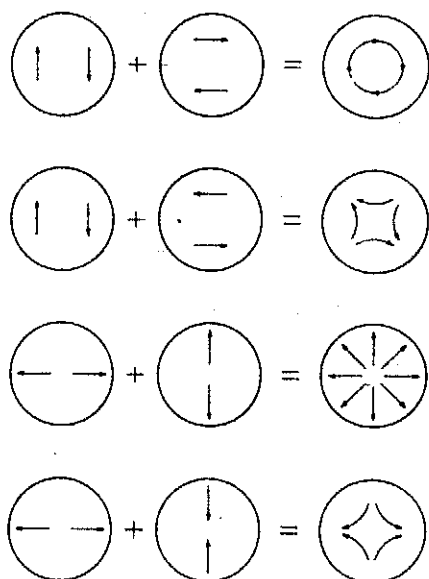


Fig. 18. Synthesis of different polarization configurations from the linearly polarized TEM_{01} mode.

laser. Linearly polarized mode configurations for square mirrors and for circular mirrors are shown in Fig. 17. By combining two orthogonally polarized modes of the same order, it is possible to synthesize other polarization configurations; this is shown in Fig. 18 for the TEM_{01} mode.

Field distributions of the resonator modes for any value of G could be obtained numerically by solving the integral equations either by the method of successive approximations [4], [24], [31] or by the method of kernel expansion [30], [32]. The former method of solution is equivalent to calculating the transient behavior of the resonator when it is excited initially with a wave of arbitrary distribution. This wave is assumed to travel back and forth between the mirrors of the resonator, undergoing changes from transit to transit and losing energy by diffraction. After many transits a quasi steady-state condition is attained where the fields for successive transits

differ only by a constant multiplicative factor. This steady-state *relative* field distribution is then an eigenfunction of the integral equations and is, in fact, the field distribution of the mode that has the lowest diffraction loss for the symmetry assumed (e.g., for even or odd symmetry in the case of infinite-strip mirrors, or for a given azimuthal mode index number l in the case of circular mirrors); the constant multiplicative factor is the eigenvalue associated with the eigenfunction and gives the diffraction loss and the phase shift of the mode. Although this simple form of the iterative method gives only the lower order solutions, it can, nevertheless, be modified to yield higher order ones [24], [35]. The method of kernel expansion, however, is capable of yielding both low-order and high-order solutions.

Figures 19 and 20 show the relative field distributions of the TEM_{00} and TEM_{01} modes for a resonator with a pair of identical, circular mirrors ($N=1$, $a_1=a_2$, $g_1=g_2=g$) as obtained by the numerical iterative method. Several curves are shown for different values of g , ranging from zero (confocal) through one (parallel-plane) to 1.2 (convex, unstable). By virtue of the equivalence property discussed in Section 4.4, the curves are also applicable to resonators with their g values reversed in sign, provided the sign of the ordinate for the phase distribution is also reversed. It is seen that the field is most concentrated near the resonator axis for $g=0$ and tends to spread out as $|g|$ increases. Therefore, the diffraction loss is expected to be the least for confocal resonators.

Figure 21 shows the relative field distributions of some of the low order modes of a Fabry-Perot resonator with (parallel-plane) circular mirrors ($N=10$, $a_1=a_2$, $g_1=g_2=1$) as obtained by a modified numerical iterative method [35]. It is interesting to note that these curves are not very smooth but have small wiggles on them, the number of which are related to the Fresnel number. These wiggles are entirely absent for the confocal resonator and appear when the resonator geometry is unstable or nearly unstable. Approximate expressions for the field distributions of the Fabry-Perot resonator modes have also been obtained by various analytical techniques [36], [37]. They are represented to first order, by sine and cosine functions for infinite-strip mirrors and by Bessel functions for circular mirrors.

For the special case of the confocal resonator ($g_1=g_2=0$), the eigenfunctions are self-reciprocal under the *finite* Fourier (infinite-strip mirrors) or Hankel (circular mirrors) transformation and exact analytical solutions exist [5], [38]–[40]. The eigenfunctions for infinite-strip mirrors are given by the prolate spheroidal wave functions and, for circular mirrors, by the generalized prolate spheroidal or hyperspheroidal wave functions. For large Fresnel number these functions can be closely approximated by Hermite-Gaussian and Laguerre-Gaussian functions which are the eigenfunction, for the beam modes.

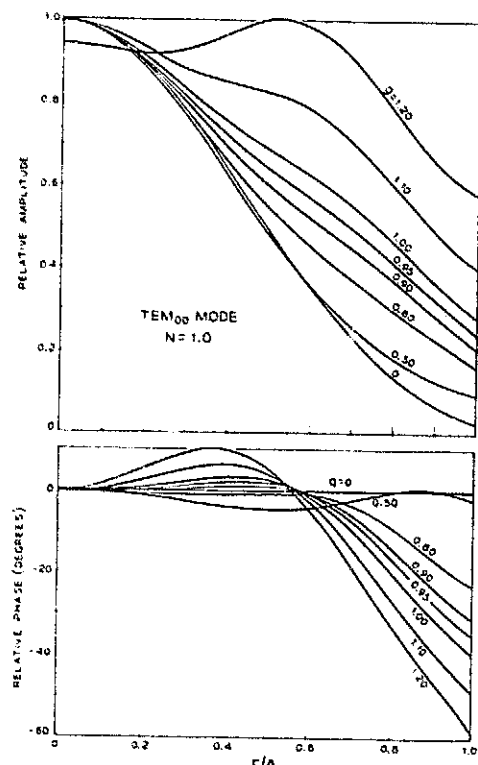


Fig. 19. Relative field distributions of the TEM_{00} mode for a resonator with circular mirrors ($N=1$).

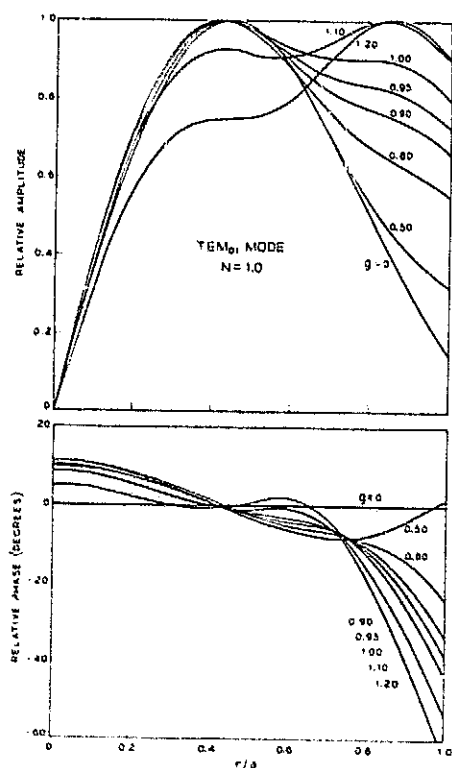


Fig. 20. Relative field distributions of the TEM_{01} mode for a resonator with circular mirrors ($N=1$).

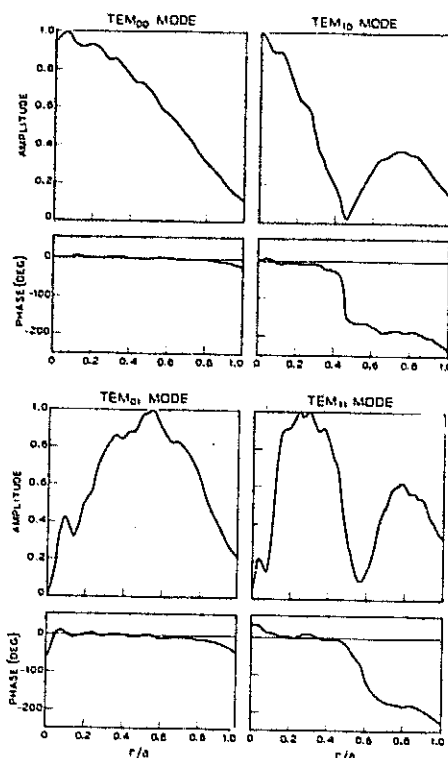


Fig. 21. Relative field distributions of four of the low order modes of a Fabry-Perot resonator with (parallel-plane) circular mirrors ($N=10$).

4.7 Diffraction Losses and Phase Shifts

The diffraction loss α and the phase shift β for a particular mode are important quantities in that they determine the Q and the resonant frequency of the resonator for that mode. The diffraction loss is given by

$$\alpha = 1 - |\gamma|^2 \quad (81)$$

which is the fractional energy lost per transit due to diffraction effects at the mirrors. The phase shift is given by

$$\beta = \text{angle of } \gamma \quad (82)$$

which is the phase shift suffered (or enjoyed) by the wave in transit from one mirror to the other, in addition to the geometrical phase shift which is given by $2\pi d/\lambda$. The eigenvalue γ in (81) and (82) is the appropriate γ for the mode under consideration. If the total resonator loss is small, the Q of the resonator can be approximated by

$$Q = \frac{2\pi d}{\lambda \alpha_t} \quad (83)$$

where α_t , the total resonator loss, includes losses due to diffraction, output coupling, absorption, scattering, and other effects. The resonant frequency ν is given by

$$\nu/\nu_0 = (q + 1) + \beta/\pi \quad (84)$$

where q , the longitudinal mode order, and ν_0 , the fundamental beat frequency, are defined in Section 3.5.

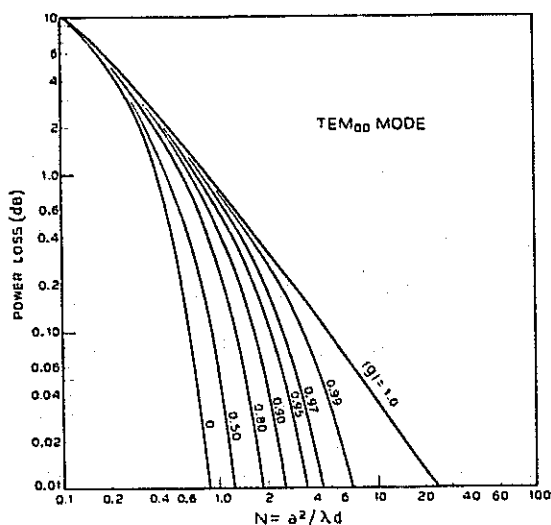


Fig. 22. Diffraction loss per transit (in decibels) for the TEM_{00} mode of a stable resonator with circular mirrors.

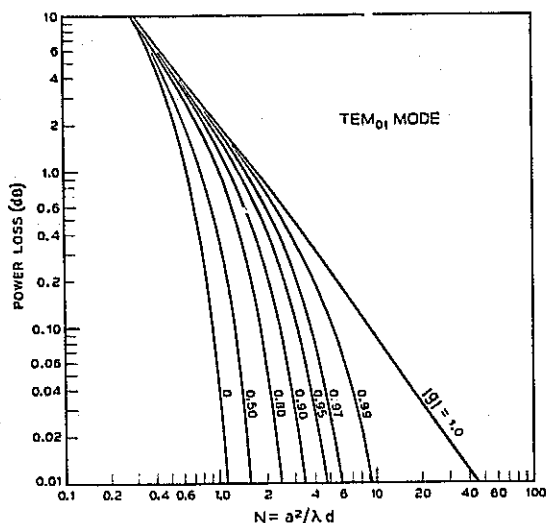


Fig. 23. Diffraction loss per transit (in decibels) for the TEM_{01} mode of a stable resonator with circular mirrors.

The diffraction losses for the two lowest order (TEM_{00} and TEM_{01}) modes of a stable resonator with a pair of identical, circular mirrors ($a_1 = a_2$, $g_1 = g_2 = g$) are given in Figs. 22 and 23 as functions of the Fresnel number N and for various values of g . The curves are obtained by solving (77) numerically using the method of successive approximations [31]. Corresponding curves for the phase shifts are shown in Figs. 24 and 25. The horizontal portions of the phase shift curves can be calculated from the formula

$$\begin{aligned} \beta &= (2p + l + 1) \arccos \sqrt{g_1 g_2} \\ &= (2p + l + 1) \arccos g, \quad \text{for } g_1 = g_2 \end{aligned} \quad (85)$$

which is equal to the phase shift for the beam modes derived in Section 3.5. It is to be noted that the loss curves are applicable to both positive and negative values of g

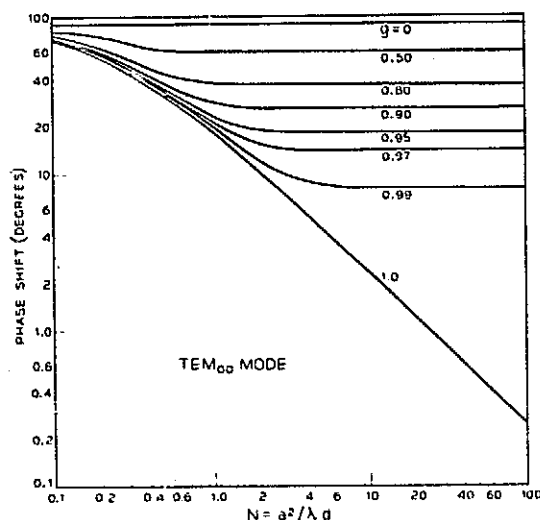


Fig. 24. Phase shift per transit for the TEM_{00} mode of a stable resonator with circular mirrors.

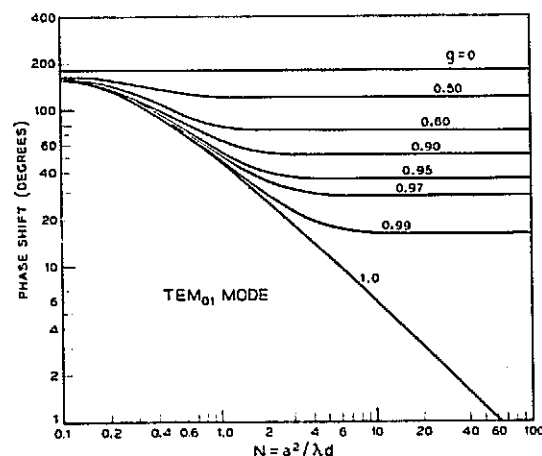


Fig. 25. Phase shift per transit for the TEM_{01} mode of a stable resonator with circular mirrors.

while the phase-shift curves are for positive g only; the phase shift for negative g is equal to 180 degrees minus that for positive g .

Analytical expressions for the diffraction loss and the phase shift have been obtained for the special cases of parallel-plane ($g=1.0$) and confocal ($g=0$) geometries when the Fresnel number is either very large (small diffraction loss) or very small (large diffraction loss) [36], [38], [39], [41], [42]. In the case of the parallel-plane resonator with circular mirrors, the approximate expressions valid for large N , as derived by Vainshtein [36], are

$$\alpha = 8\kappa_{pt} \frac{\delta(M + \delta)}{[(M + \delta)^2 + \delta^2]^2} \quad (86)$$

$$\beta = \left(\frac{M}{4\delta}\right) \alpha \quad (87)$$

where $\delta=0.824$, $M=\sqrt{8\pi N}$, and κ_{pl} is the $(p+1)$ th zero of the Bessel function of order l . For the confocal resonator with circular mirrors, the corresponding expressions are [39]

$$\alpha = \frac{2\pi(8\pi N)^{2p+l+1}e^{-4\pi N}}{p!(p+l+1)!} \left[1 + O\left(\frac{1}{2\pi N}\right) \right] \quad (88)$$

$$\beta = (2p+l+1) \frac{\pi}{2} \quad (89)$$

Similar expressions exist for resonators with infinite-strip or rectangular mirrors [36], [39]. The agreement between the values obtained from the above formulas and those from numerical methods is excellent.

The loss of the lowest order (TEM₀₀) mode of an unstable resonator is, to first order, independent of the mirror size or shape. The formula for the loss, which is based on geometrical optics, is [12]

$$\alpha = 1 \pm \frac{1 - \sqrt{1 - (g_1 g_2)^{-1}}}{1 + \sqrt{1 - (g_1 g_2)^{-1}}} \quad (90)$$

where the plus sign in front of the fraction applies for g values lying in the first and third quadrants of the stability diagram, and the minus sign applies in the other two quadrants. Loss curves (plotted vs. N) obtained by solving the integral equations numerically have a ripply behavior which is attributable to diffraction effects [24], [43]. However, the average values agree well with those obtained from (90).

5. CONCLUDING REMARKS

Space limitations made it necessary to concentrate the discussion of this article on the basic aspects of laser beams and resonators. It was not possible to include such interesting topics as perturbations of resonators, resonators with tilted mirrors, or to consider in detail the effect of nonlinear, saturating host media. Also omitted was a discussion of various resonator structures other than those formed of spherical mirrors, e.g., resonators with corner cube reflectors, resonators with output holes, or fiber resonators. Another important, but omitted, field is that of mode selection where much research work is currently in progress. A brief survey of some of these topics is given in [44].

REFERENCES

- [1] R. H. Dicke, "Molecular amplification and generation systems and methods," U. S. Patent 2 851 652, September 9, 1958.
- [2] A. M. Prokhorov, "Molecular amplifier and generator for submillimeter waves," *JETP (USSR)*, vol. 34, pp. 1658-1659, June 1958; *Sov. Phys. JETP*, vol. 7, pp. 1140-1141, December 1958.
- [3] A. L. Schawlow and C. H. Townes, "Infrared and optical masers," *Phys. Rev.*, vol. 29, pp. 1940-1949, December 1958.
- [4] A. G. Fox and T. Li, "Resonant modes in an optical maser," *Proc. IRE (Correspondence)*, vol. 48, pp. 1904-1905, November 1960; "Resonant modes in a maser interferometer," *Bell Sys. Tech. J.*, vol. 40, pp. 453-488, March 1961.
- [5] G. D. Boyd and J. P. Gordon, "Confocal multimode resonator for millimeter through optical wavelength masers," *Bell Sys. Tech. J.*, vol. 40, pp. 489-508, March 1961.
- [6] G. D. Boyd and H. Kogelnik, "Generalized confocal resonator theory," *Bell Sys. Tech. J.*, vol. 41, pp. 1247-1269, July 1962.
- [7] G. Goubau and F. Schwing, "On the guided propagation of electromagnetic wave beams," *IRE Trans. on Antennas and Propagation*, vol. AP-9, pp. 248-256, May 1961.
- [8] J. R. Pierce, "Modes in sequences of lenses," *Proc. Nat'l Acad. Sci.*, vol. 47, pp. 1808-1813, November 1961.
- [9] G. Goubau, "Optical relations for coherent wave beams," in *Electromagnetic Theory and Antennas*, New York: Macmillan, 1963, pp. 907-918.
- [10] H. Kogelnik, "Imaging of optical mode resonators with internal lenses," *Bell Sys. Tech. J.*, vol. 44, pp. 455-494, March 1965.
- [11] —, "On the propagation of Gaussian beams of light through lenslike media including those with a loss or gain variation," *Appl. Opt.*, vol. 4, pp. 1562-1569, December 1965.
- [12] A. E. Siegman, "Unstable optical resonators for laser applications," *Proc. IEEE*, vol. 53, pp. 277-287, March 1965.
- [13] W. Brower, *Matrix Methods in Optical Instrument Design*, New York: Benjamin, 1964. E. L. O'Neill, *Introduction to Statistical Optics*, Reading, Mass.: Addison-Wesley, 1963.
- [14] M. Bertolotti, "Matrix representation of geometrical properties of laser cavities," *Nuovo Cimento*, vol. 32, pp. 1242-1257, June 1964. V. P. Bykov and L. A. Vainshtein, "Geometrical optics of open resonators," *JETP (USSR)*, vol. 47, pp. 508-517, August 1964. B. Macke, "Laser cavities in geometrical optics approximation," *J. Phys. (Paris)*, vol. 26, pp. 104A-112A, March 1965. W. K. Kahn, "Geometric optical derivation of formula for the variation of the spot size in a spherical mirror resonator," *Appl. Opt.*, vol. 4, pp. 758-759, June 1965.
- [15] J. R. Pierce, *Theory and Design of Electron Beams*, New York: Van Nostrand, 1954, p. 194.
- [16] H. Kogelnik and W. W. Rigrod, "Visual display of isolated optical-resonator modes," *Proc. IRE (Correspondence)*, vol. 50, p. 220, February 1962.
- [17] G. A. Deschamps and P. E. Mast, "Beam tracing and applications," in *Proc. Symposium on Quasi-Optics*, New York: Polytechnic Press, 1964, pp. 379-395.
- [18] S. A. Collins, "Analysis of optical resonators involving focusing elements," *Appl. Opt.*, vol. 3, pp. 1263-1275, November 1964.
- [19] T. Li, "Dual forms of the Gaussian beam chart," *Appl. Opt.*, vol. 3, pp. 1315-1317, November 1964.
- [20] T. S. Chu, "Geometrical representation of Gaussian beam propagation," *Bell Sys. Tech. J.*, vol. 45, pp. 287-299, February 1966.
- [21] J. P. Gordon, "A circle diagram for optical resonators," *Bell Sys. Tech. J.*, vol. 43, pp. 1826-1827, July 1964. M. J. Offerhaus, "Geometry of the radiation field for a laser interferometer," *Philips Res. Rept.*, vol. 19, pp. 520-523, December 1964.
- [22] H. Statz and C. L. Tang, "Problem of mode deformation in optical masers," *J. Appl. Phys.*, vol. 36, pp. 1816-1819, June 1965.
- [23] A. G. Fox and T. Li, "Effect of gain saturation on the oscillating modes of optical masers," *IEEE J. of Quantum Electronics*, vol. QE-2, p. 1xii, April 1966.
- [24] —, "Modes in a maser interferometer with curved and tilted mirrors," *Proc. IEEE*, vol. 51, pp. 80-89, January 1963.
- [25] F. B. Hildebrand, *Methods of Applied Mathematics*, Englewood Cliffs, N. J.: Prentice Hall, 1952, pp. 412-413.
- [26] D. J. Newman and S. P. Morgan, "Existence of eigenvalues of a class of integral equations arising in laser theory," *Bell Sys. Tech. J.*, vol. 43, pp. 113-126, January 1964.
- [27] J. A. Cochran, "The existence of eigenvalues for the integral equations of laser theory," *Bell Sys. Tech. J.*, vol. 44, pp. 77-88, January 1965.
- [28] H. Hochstadt, "On the eigenvalue of a class of integral equations arising in laser theory," *SIAM Rev.*, vol. 8, pp. 62-65, January 1966.
- [29] D. Gloge, "Calculations of Fabry-Perot laser resonators by scattering matrices," *Arch. Elect. Uebertrag.*, vol. 18, pp. 197-203, March 1964.
- [30] W. Streifer, "Optical resonator modes—rectangular reflectors of spherical curvature," *J. Opt. Soc. Am.*, vol. 55, pp. 868-877, July 1965.
- [31] T. Li, "Diffraction loss and selection of modes in maser resonators with circular mirrors," *Bell Sys. Tech. J.*, vol. 44, pp. 917-932, May-June, 1965.
- [32] J. C. Heurtley and W. Streifer, "Optical resonator modes—

- circular reflectors of spherical curvature," *J. Opt. Soc. Am.*, vol. 55, pp. 1472-1479, November 1965.
- [33] J. P. Gordon and H. Kogelnik, "Equivalence relations among spherical mirror optical resonators," *Bell Sys. Tech. J.*, vol. 43, pp. 2873-2886, November 1964.
 - [34] F. Schwing, "Reiterative wave beams of rectangular symmetry," *Arch. Elect. Übertrag.*, vol. 15, pp. 555-564, December 1961.
 - [35] A. G. Fox and T. Li, to be published.
 - [36] L. A. Vainshtein, "Open resonators for lasers," *JETP (USSR)*, vol. 44, pp. 1050-1067, March 1963; *Sov. Phys. JETP*, vol. 17, pp. 709-719, September 1963.
 - [37] S. R. Barone, "Resonances of the Fabry-Perot laser," *J. Appl. Phys.*, vol. 34, pp. 831-843, April 1963.
 - [38] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—I," *Bell Sys. Tech. J.*, vol. 40, pp. 43-64, January 1961.
 - [39] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—IV: Extensions to many dimensions; generalized prolate spheroidal functions," *Bell Sys. Tech. J.*, vol. 43, pp. 3009-3057, November 1964.
 - [40] J. C. Heurtley, "Hyperspheroidal functions—optical resonators with circular mirrors," in *Proc. Symposium on Quasi-Optics*. New York: Polytechnic Press, 1964, pp. 367-375.
 - [41] S. R. Barone and M. C. Newstein, "Fabry-Perot resonances at small Fresnel numbers," *Appl. Opt.*, vol. 3, p. 1194, October 1964.
 - [42] L. Bergstein and H. Schachter, "Resonant modes of optic cavities of small Fresnel numbers," *J. Opt. Soc. Am.*, vol. 55, pp. 1226-1233, October 1965.
 - [43] A. G. Fox and T. Li, "Modes in a maser interferometer with curved mirrors," in *Proc. Third International Congress on Quantum Electronics*. New York: Columbia University Press, 1964, pp. 1263-1270.
 - [44] H. Kogelnik, "Modes in optical resonators," in *Lasers*, A. K. Levine, Ed. New York: Dekker, 1966.

APPENDIX D. PROPERTIES OF MATERIALS FROM 100 TO 200 GHz

1. INTRODUCTION

Surface resistivities of metals and dissipation factors (loss tangents) of dielectrics cause ohmic losses in the radiometer signal channel. Metallic losses are predictable in both waveguides and quasi-optical assemblies because surface resistivities of metals and behavior are well known. However, dielectric data in the frequency region from 100 to 200 GHz are incomplete and spotty. This study shows the need for a dielectric materials measurement program.

Section 2 lists the reflective losses encountered from various types of metals. The theoretical values are more optimistic than measured values by a typical factor of 3. Section 3 lists dielectric data obtained from an exhaustive survey with all sources referenced. In cases where 100 to 200 GHz data are not available, data from other frequencies are listed. Section 4 discusses two effective methods of reducing reflective losses from dielectric surfaces.

2. REFLECTIVE LOSSES FROM METALS

Reflective losses from metals can be shown to be given by

$$R_r = \frac{2\omega\delta}{c}$$

where

R_r = fraction of energy absorbed by metal surface

δ = skin depth of metal

ω = frequency

c = velocity of light*

Skin depth δ is defined to be the distance an electromagnetic wave field penetrates into a metal**

$$\delta = \left\{ \omega \left[\frac{\mu\epsilon}{2} \left(\sqrt{1 + \left(\frac{\sigma}{\mu\epsilon} \right)^2} - 1 \right) \right]^{1/2} \right\}^{-1}$$

where

μ = magnetic permeability

ϵ = dielectric permittivity

σ = surface resistivity of metal

If the material is a good conductor, $\sigma \gg \omega\epsilon$ and

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}}$$

*G. P. Harnwell, Principles of Electricity and Electromagnetism, McGraw Hill, New York, 1949, p. 587.

**Fink, Electronic Engineer's Handbook.

This approximation holds for frequencies beyond 1000 GHz for metals of typical resistivities.

Table 1 lists surface resistivities, skin depths, and reflective losses at 2, 100, and 200 GHz for ten common metals used in microwave and millimeter wave circuits. It should be observed that theoretical reflective losses from surfaces of copper and silver are the order of 0.1 percent. In practice, these are approximately 0.3 percent at 200 GHz.

TABLE 1. REFLECTIVE LOSSES OF COMMON METALS

Metal	Resistivity ($\Omega\text{-m}$) $\times 10^{-8}$	Skin Depth, μm			Reflective Loss		
		2 GHz	100 GHz	200 GHz	2 GHz	100 GHz	200 GHz
Aluminum	2.62	1.82 [†]	0.258	0.182	0.000152	0.0012	0.0015
Beryllium	4.57	2.41	0.340	0.241	0.000202	0.0014	0.0020
Brass (66 cu 34 zn)	3.9	2.2	0.31	0.22	0.000185	0.0013	0.0018
Chromium	2.6	1.8	0.26	0.18	0.000151	0.0011	0.00151
Copper	1.72	1.48 [†]	0.209	0.148	0.000124	0.0008	0.00124
Gold	2.44	1.76 [†]	0.249	0.176	0.000148	0.0010	0.00147
Nickel	6.9	0.42 [†]	0.059*	0.042*			
Platinum	10.5	3.65	0.516	0.365	0.000307	0.002	0.0031
Silver	1.62	1.43 [†]	0.203	0.143	0.000120	0.0008	0.0012
Tin	11.4	3.80	0.537	0.380	0.000319	0.0022	0.0032

* Assumes $\mu_r = 50$.

[†] Verified by comparison with published values; see Fink Electronics Engineers' Handbook, pg. 6-4.

3. DIELECTRIC DISSIPATION FACTOR (LOSS TANGENT)

Equations and Definitions

The complex relative permittivity $\hat{\epsilon}$ and complex index of refraction $\hat{n}$ are related by

$$\hat{\epsilon} = (\hat{n})^2$$

where

$$\hat{\epsilon} = \epsilon' - j\epsilon''$$

and

$$\hat{n} = n - j \frac{c\alpha}{4\pi\nu} = n(1 - jK)$$

c = speed of light = 3×10^8 m/sec

K = extinction coefficient = $c\alpha/4\pi\nu n$

α = absorption coefficient per meter

n = real index of refraction

Now

$$\hat{\epsilon} = n^2 - \left(\frac{c\alpha}{4\pi\nu}\right)^2 - \frac{jnc\alpha}{2\pi\nu}$$

equating components gives

$$\epsilon' = n^2 - \left(\frac{c\alpha}{4\pi\nu}\right)^2$$

and

$$\epsilon'' = \frac{cn\alpha}{2\pi v}$$

$$\tan \delta = \frac{\epsilon''}{\epsilon'} = \frac{8\pi v n c \alpha}{(4\pi v n)^2 - (c\alpha)^2}$$

δ = loss angle in μrad

When

$$\alpha \ll \frac{4\pi v n}{c}$$

then

$$\epsilon' \approx n^2$$

and

$$\tan \delta \approx \frac{c\alpha}{2\pi v n}$$

} very useful approximations

In general, these approximations are made whenever $\alpha \ll v/c$ and $\mu' = \mu_0$ (magnetization can be neglected for organic polymers). Then if absorption is assumed to be the only source of attenuation, we may write an expression for transmission τ

$$\tau = \exp[-\alpha \ell]$$

where ℓ is the thickness of dielectric material in question.

Table 2 lists the dielectric properties of materials.

TABLE 2. DIELECTRIC PROPERTIES OF MATERIALS, GHz

Material	Dielectric Constant ϵ'	Dielectric Loss ϵ''	Loss Tangent $\epsilon''/\epsilon' = \tan \delta$	Absorption Coefficient ⁽¹⁾ (neper-cm ⁻¹)	Reference
Polyethylene	2.24 25	0.0005 25	0.00021 25	0.0017 25	i
	2.31 143				b
	2.31 343	0.00048 300	0.00021 300	0.02 300	b and j
	2.33 850	0.0009 850	0.0004 850	0.1087 850	h
Polypropylene	2.25 35				e
	2.25* 120	0.0012 120	0.00053 120	0.02 120	f
Polystyrene	2.54 25	0.003 25	0.0012 25	0.010 25	i
	2.52 140	0.005 140	0.002 140	0.0924 140	a
	2.53 210	0.00316 120	0.0013 120	0.05 120	a and d
	2.57 343	0.0063 168	0.0025 168	0.14 168	b and d
	2.52 850	0.00227 850	0.0009 850	0.25 850	h
Rexolite (1422) (cross linked polystyrene)	2.54 10	0.00127 10	0.0005 10	0.0017 10	i
	2.47 140	0.00494 140	0.002 140	0.0922 140	a
	2.50 210				a
	2.54 343				b
T.P.X. (basically a poly 4-methyl 1-pentene)	2.126 35	0.001 35	0.00048 35	0.0051 35	g
	2.15* 120	0.0016 120	0.00076 120	0.028 120	f
Teflon (treated PTFE, polytetrafluoroethylene)	2.08 25	0.00125 25	0.0006 25	0.0045 25	i
	2.058 35	0.00062 50	0.0003 50	0.0045 50	e and k
	2.039 60-90				e
	2.05 140	0.00615 140	0.003 140	0.126 140	a
	2.08 210				a
	2.07 343				b
PTFE (unsintered)	2.042 850	0.00143 850	0.0007 850	0.95 850	h
	1.950 35				g
	1.952 35	0.0001 35	0.00005 35	0.00049 35	i
	2.052 300	0.0002 300	0.0001 300	0.009 ⁽²⁾ 300	j
RTFE ⁽³⁾ (sintered)	1.933 300	0.00084 300	0.000435 300	0.38 300	j
Paraffin	2.2 25	0.00066 25	<0.0003 25	<0.0023 25	i
	2.19 120	0.00595 120	0.0027 120	0.10 120	d
	2.2 168	0.00296 168	0.00135 168	0.07 168	d
	2.3 2-300		0.001		c
Lucite	2.57 25	0.0082 25	0.0032 25	0.027 25	i
	2.56 140	0.038 140	0.015 140	0.696 140	a
	2.58 210				a
Plexiglas	2.59 25	0.017 25	0.0067 25	0.056 25	i
	2.59 120	0.019 120	0.0075 120	0.30 120	d
	2.60 143	0.023 168	0.0089 168	0.49 168	b and d
	2.61 343				b
Mylar (polyethylene terephthalate)	3.35 140	0.0335 140	0.01 140	0.537 140	a
	3.145 55	0.0138 55	0.0044 55	0.09 55	m
Fused Silica (85% density slip cast)	3.19 94	0.0085 94	0.0026 94	0.0923 94	h
Fused Silica (spectrosil)	3.78 10	0.00084 10	0.00017 10	0.0007 10	i
	3.85 2-300	0.00039	0.0001	0.003 75 ⁽⁶⁾	c
	3.82 60-90				e
Glass (Corning 7070)	3.9 25	0.0121 25	0.0031 25	0.032 25	i
	4.0 2-300	0.0096	0.0024	—	c
Quartz (crystal-anisotropic)	4.43 ⁽⁴⁾ 35	0.00014 35	0.000031 35	0.00048 35	g
	4.45 ⁽⁴⁾ 2-300		0.0002	—	c
	4.64 ⁽⁵⁾ 35				
Ethyl Cellulose	2.65 25	0.0795 25	0.03 25	0.256 25	i
	3.71 140	0.371 140	0.1 140	5.65 140	a

Note: Temperature is usually at or near 25°C or 300°K.

* Assumed value necessary to calculate absorption.

1) Absorption coefficient per unit length is without regard to effects at the medium boundaries. The neper here, as defined in submillimetre optics, is only half the magnitude of the neper used in E.E.; therefore, 1 Neper = 4.343 dB is the conversion.

2) This is an upper bound corresponding to the limit of the particular experimental technique used in reference j.

3) "Sintered" PTFE is obtained by stretching a sample at 250°C and allowing it to cool under tension.

4) Ordinary ray ϵ_0 .5) Extraordinary ϵ_0 .

6) Arbitrarily chosen frequency between 2 and 300 GHz.

Table 2 (continued)

REFERENCES TO TABLE 2

- a. F. Sobel, F. L. Wentworth, "Quasi-optical Surface Waveguide and Other Components for 100 to 300 Gc Region," IRE TRANS on Microwave Theory and Techniques, Volume MTT 9, pp. 512-518, November 1961.
- b. R. Fellers, "Measurements in the Millimeter to Micron Range," Proceedings of the IEEE, vol. 55, pp. 1003-1014, June 1967.
- c. Harvey, Microwave Engineering, pp. 252-253, Academic Press, 1963.
- d. W. L. Brooks, et al., "Absorption of Millimeter Waves in Dielectric Solids," Journal of the Optical Society of America, Vol. 43, pp. 1191-1194, December 1953.
- e. R. G. Jones, "Millimeter Wave Dielectric Measurement Using Open Resonators," High Frequency Dielectric Measurement, pp. 78-82 IPC Science and Technology Press Ltd., 1972.
- f. G. W. Chantry, "Optical Materials for the Submillimeter Wave Band," High Frequency Dielectric Measurement, pp. 117-121, IPC Science and Technology Press Ltd., 1972.
- g. R. G. Jones, "Dielectric Measurements at MM Wavelengths Using Open Resonators," Dielectric Materials Measurements and Applications, pp. 141-144, IEE Conference Publication No. 129, 1975.
- h. K. H. Breeden, et al., "Complex Permittivity Measurements at Millimeter Wavelengths," Dielectric Materials Measurements and Applications, pp. 50-53, IEE Conference Publication No. 67, 1970.
- i. A. R. Von Hippel, Dielectric Materials and Applications, John Wiley and Sons, Inc., New York, 1956.
- j. J. Fleming and G. Chantry, "Accurate Radiometric Measurements on Low-Loss Polymers at Submillimeter Wavelengths," IEEE Trans on Instrumentation and Measurement, pp. 473-478, Vol. IM-23, No. 4, December 1974.
- k. W. Culshaw and M. V. Anderson, "Measurement of Permittivity and Dielectric Loss With a MM-Wave Fabry-Perot Interferometer," IEE Proceedings, Vol. 109, Part B, Supp. 23, pp. 820-826, 1961.
- l. R. J. Cook, R. G. Jones, and C. B. Rosenberg, "Comparison of Cavity and Open Resonator Measurements of Permittivity and Loss Angle," IEEE Trans on Instrumentation and Measurement, Vol. IM-23, No. 4, pp. 438-442, December 1974.
- m. G. E. Conklin, "Measurement of the Dielectric Loss Tangent of Isotropic Films at Millimeter Wavelengths," Review of Scientific Instruments, pp. 1347-1349, Vol. 36, No. 9, September 1965.
- n. G. W. Chantry, et al., "Far Infrared and Millimeter Wave Absorption Spectra of Some Low-Loss Polymers," Chemical Physics Letters, pp. 473-477, Vol. 10, No. 4, 15 August 1971.
- o. M. N. Afshar and G. W. Chantry, "Precise Dielectric Measurements of Low-Loss Materials at Millimeter and Submillimeter Wavelengths," IEEE Transactions on Microwave Theory and Technique, pp. 509-511, Vol. MTT-25, No. 6, June 1977.
- p. M. E. A. Cudby, et al., "The Vibrational Spectrum of Polytetrafluorethylene in the Submillimeter Region," Dielectric Materials Measurements and Applications, pp. 120-125, IEE Conference, No. 129, 1975.
- q. J. Chamberlain, "Submillimeter-Wave Techniques," High Frequency Dielectric Measurement, p. 105, IPC Science and Technology Press Ltd., 1972.

Note: Reference c provided loss tangents taken as "average results" from about 2 to 300 GHz.

4. DIELECTRIC REFLECTION LOSSES

Reflective losses from a dielectric surface are determined by

$$L_r = \left(\frac{n - 1}{n + 1} \right)^2$$

where n is the refractive index of the material or $\sqrt{\epsilon}$. Typical losses for an air-quartz surface are the order of 1 dB. A quartz lens having two surfaces will have loss contributions from both surfaces or 2 dB.

Reflective losses can be minimized by two methods (see Figure D-1). A common method in optics is a $\lambda/4$ coating on the surface of proper refractive index. For example, if a fused silica lens $n_s = 1.962$ and teflon $n_t = 1.43$, the reflective losses of the quartz can be reduced by coating the quartz with a layer of teflon. The reflective loss with coating is then

$$L_r = \frac{(R_{10} + R_{21})^2 - 4R_{10}R_{21}}{1 + R_{10}R_{21}}$$

where

$$R_{10} = \frac{n_t - 1}{n_t + 1}$$

and

$$R_{21} = \frac{n_s - n_t}{n_s + n_t}$$

which reduces the loss to 0.019, or 0.08 dB per surface.

Another method of eliminating surface reflections is to machine grooves in the optical surface to simulate the quarter wave spacing. This technique applies especially to materials like Rexolite which machine easily, but can also be achieved with Teflon. The technique is to adjust the ratio of

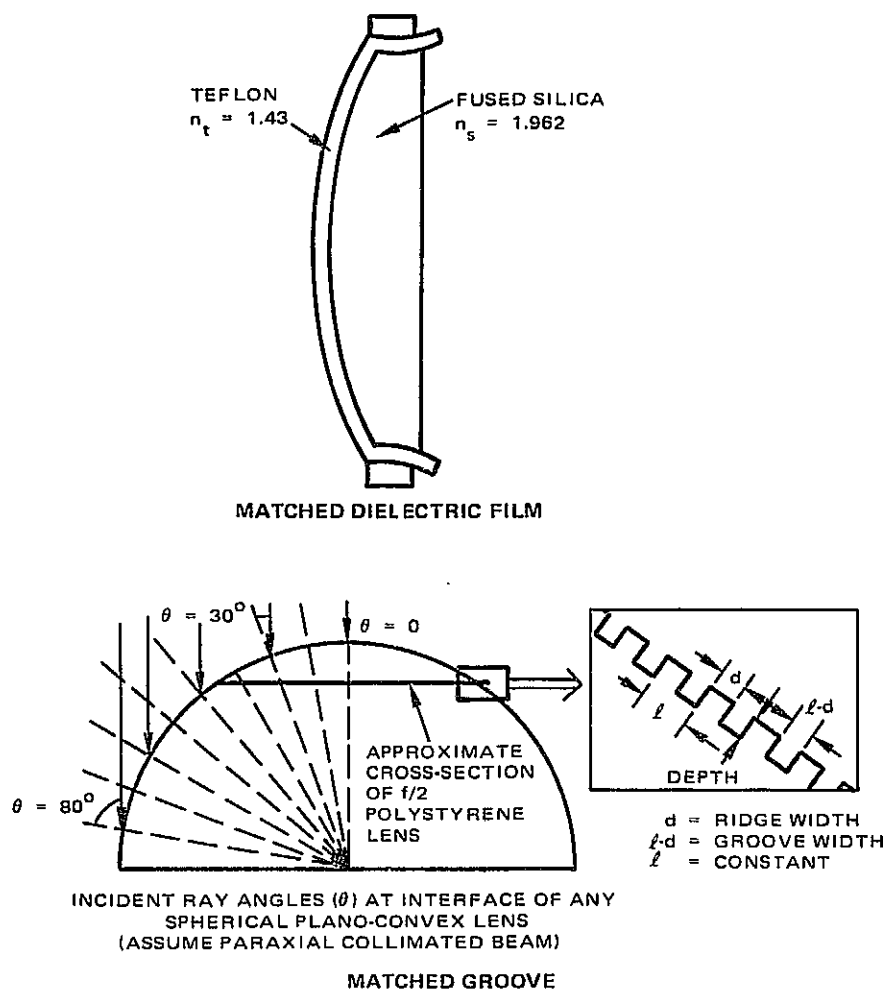


FIGURE D-1. REDUCTION OF REFLECTION LOSSES BY DIELECTRIC FILM AND GROOVE MATCHINGS

groove width to groove spacing to achieve the correct "effective" matching index of refraction, and to adjust the depth of the grooves to obtain the quarter wave spacing.*

*S. B. Cohn and T. Morita, "Measured Performance of Dielectric Lenses," IRE Transactions on Antennas and Propagation, January 1956, pp. 31-33.